

Solutions de recherche pour les entreprises

Intégration de la recherche à votre architecture IT

Les solutions de recherche pour entreprises peuvent être **bénéfiques en termes de productivité et de performance commerciale**, à condition que leur mise en œuvre soit efficace.

INTRODUCTION : RÔLE DE L'ARCHITECTURE

Les solutions de recherche pour entreprises apparaissent désormais comme des outils puissants, permettant aux entreprises d'atteindre un objectif universel : utiliser l'information de manière optimale.

Les solutions de recherche multiniveaux visent à augmenter la productivité et l'efficacité des organisations en analysant les informations se trouvant sur les postes de travail, au niveau des différents services de l'entreprise et sur Internet. Ces solutions permettent de surmonter les principales difficultés que rencontrent aujourd'hui les organisations : stockage redondant de données identiques, méthodes de récupération diverses et absence de définitions sémantiques cohérentes au sein de l'entreprise. Elles peuvent également satisfaire les exigences croissantes des employés qui attendent du système de leur entreprise qu'il soit aussi facile à explorer et aussi convivial que les logiciels et services Internet.

Afin de bénéficier pleinement des avantages d'une fonction de recherche en entreprise, il est nécessaire d'intégrer la solution de recherche de manière effective aux systèmes technologiques de l'organisation. BearingPoint a identifié les principes fondamentaux de la fonction de recherche pour les entreprises qui lui ont permis de développer une architecture modèle assurant un déploiement réussi.

LES PRINCIPES DE LA RECHERCHE EN ENTREPRISE

Plusieurs principes généraux régissent le développement architectural d'une solution de recherche :

Simplicité : En principe, une fonction de recherche portant sur l'ensemble d'une entreprise doit fonctionner tout en préservant la simplicité des systèmes et des solutions. En présence d'un système complexe, la productivité des développeurs de logiciels est bridée ; la conception, la construction et les tests de produits deviennent délicats ; les problèmes de sécurité se transforment en défis, provoquant un sentiment de frustration chez les utilisateurs et les administrateurs. Les technologies informatiques et de communication ont évolué à tel point qu'adopter un modèle de services est maintenant une solution viable. Les logiciels optimisés pour les services et le développement de solutions de recherche légères répondent aux attentes des utilisateurs en fournissant un système facile d'utilisation et convivial qui "fonctionne tout simplement".

Fonctionnalité : L'architecture d'une solution de recherche doit posséder certaines caractéristiques essentielles. Elle doit notamment être (par ordre d'importance) : utilisable, fiable, sécurisée, évolutive (en termes de nombre d'utilisateurs, de volumes et de services de messagerie) et modulable afin d'ajouter de nouveaux services proposés dans le cadre de la solution de recherche.

SOMMAIRE :

INTRODUCTION : RÔLE DE L'ARCHITECTURE	1
LES PRINCIPES DE LA RECHERCHE EN ENTREPRISE	1
CONSIGNES À RESPECTER DANS LE DÉVELOPPEMENT DE SOLUTIONS	2
ARCHITECTURE DE LA SOLUTION DE RECHERCHE	3
Architecture conceptuelle	4
Exemple d'architecture pour solution de recherche	4
SOLUTION DE RECHERCHE, UNE OPPORTUNITÉ À SAISIR	8

Fiabilité : La plateforme d'une solution de recherche nécessite une infrastructure solide. Par conséquent, l'utilisation de technologies de pointe est à éviter. De manière générale, la technologie de son architecture centrale doit :

- fournir une transportabilité maximale dans des environnements divers ;
- avoir été testée et éprouvée dans des conditions d'utilisation comparables.

Point critique du système : La solution de recherche d'une entreprise doit être considérée comme une composante essentielle de l'activité de la société, offrant une sécurité et une évolutivité élevées ainsi qu'une disponibilité 24 heures sur 24. Son architecture doit être établie dans le respect de ces principes fondateurs.

Accessibilité : L'architecture d'une solution de recherche doit permettre un accès étendu afin d'étendre sa capacité. Cette dernière doit être intégrée au moment de la conception, car elle ne peut pas être greffée au système ultérieurement.

Flexibilité: Le concept de services flexibles est déterminant pour la réussite d'une solution de recherche non pas du point de vue du service proposé à l'utilisateur final, mais en termes d'unités de fonctionnement distinctes et accessibles sur la plateforme de recherche par le biais d'interfaces définies.

CONSIGNES À RESPECTER DANS LE DÉVELOPPEMENT DE SOLUTIONS

Selon les principes énoncés ci-dessus, l'architecture d'une solution de recherche doit respecter les règles suivantes :

Normes et technologies ouvertes et interfaces communes : Toute solution de recherche doit reposer sur des normes ouvertes assurant une indépendance technologique grâce à l'utilisation de normes telles que J2EE, .NET et services Web. Cela apporte une flexibilité optimale, permet d'éviter les modifications d'interface lorsque l'interface d'une autre application change, et dissimule la connectivité complexe des interfaces de programmation d'applications (API) entre des systèmes et des applications associés.

Architecture modulaire : Les solutions de recherche sont basées sur une architecture multinationnelle organisée en composants. Cette approche fait abstraction des couches du système et confère à chaque composant une indépendance et un fonctionnement individuel et logique. Le couplage d'une telle architecture doit être flexible et extrêmement précis.

Interopérabilité : L'utilisation des langages XML (extensible markup language) et XSLT (extensible style sheet language transformation) crée une certaine interopérabilité permettant de dissocier le domaine de la présentation de la logique d'arrière-plan. Le fait de dissocier présentation et interface de présentation dans deux sous-couches distinctes permet d'isoler la logique de présentation. L'interopérabilité repose sur l'indépendance des périphériques, des plateformes et des technologies établie par l'architecture.

Gestion des métadonnées : Cette fonction permet de gérer la création, le stockage, l'accès, l'évolution et la maintenance des contenus en métadonnées, notamment dans le cadre de la gestion et de la synchronisation de métadonnées partagées. Dans une société, la gestion de métadonnées permet d'effectuer les opérations suivantes :

- Automatiser l'attribution de métadonnées à des contenus existants. L'efficacité de la recherche est ainsi accrue, car les ensembles de contenus sont détectés, quel que soit leur emplacement.
- Développer la capacité à associer des métadonnées à des ensembles de contenus aussi rapidement que possible à la suite de leur création.
- Établir une administration précise de l'activité en matière de développement et de gestion quotidienne des classifications et des listes d'autorités.
- Mettre en œuvre des processus d'actualisation des classifications et de suivi des associations de métadonnées.
- Obtenir diverses présentations d'un même contenu, notamment grâce au tri par rôle, heure et processus.
- Intégrer les sources de contenu : la solution de recherche doit posséder une architecture ouverte afin de prendre en charge l'intégration de plusieurs sources de contenus. Les normes ouvertes s'avèrent nécessaires pour faciliter le fonctionnement de la plateforme, en particulier pour assurer l'interopérabilité de plusieurs sources de contenu.

Les solutions de recherche multiniveaux permettent de surmonter les principales difficultés que rencontrent aujourd'hui les organisations.

Sécurité: Une solution de recherche pour entreprise doit tenir compte des principes de sécurité appliqués aux différentes sources de contenus de l'entreprise et les développer. Les règles relatives à l'authentification, aux autorisations, au suivi et à la gestion des identités doivent être appliquées et assurer que les informations à accès restreint ne soient accessibles qu'aux utilisateurs autorisés. La solution de recherche doit être conforme à la politique de sécurité de l'entreprise et aux exigences réglementaires. Dans certains cas, il est nécessaire d'établir une sécurité au niveau des documents, afin que l'utilisation de contenus non structurés soit effective.

Infrastructure réseau : Le déploiement d'une solution doit tenir compte de l'impact sur l'infrastructure réseau et de l'organisation à prendre en charge. Dans le cas d'une recherche d'entreprise, cela inclut l'emplacement des données à explorer. La fonction de recherche, la sécurité et les performances globales sont également des facteurs à prendre en considération si des personnes ou organisations extérieures font usage de la solution.

Les questions à poser en termes de conception générale du réseau sont notamment :

- Quel est le volume de trafic généré par les recherches effectuées sur le réseau de l'entreprise et les segments associés ? Le nombre de sources de données explorées, leur emplacement et le type d'infrastructure déployé permettront de définir la conception à mettre en œuvre.
- Où les utilisateurs se situent-ils par rapport à la solution de recherche : dans le même bâtiment, à une autre adresse ou sur un réseau étendu (WAN) ?
- Une classification du trafic a-t-elle été établie ? La classification du trafic est une question importante dans un environnement qui pratique la duplication de données d'espaces de stockage en réseau (SAN), tels que Symmetrix Remote Data Facility (SRDF). La société doit tenir compte de l'impact de la recherche sur le réseau, mettre en place des règles de segmentation réseau et une qualité de service appropriées.

- Quel est le type de contrôle d'accès disponible sur le réseau ? Dans de nombreux cas, le déploiement constitue une phase cruciale, car la solution de recherche doit permettre d'afficher et de diffuser des données protégées ou confidentielles. La segmentation réseau, la conception du réseau local virtuel (VLAN), la configuration du pare-feu et la connexion au réseau doivent être prises en compte.
- Quel type de résilience système est nécessaire ? Pour une entreprise, la solution de recherche est souvent considérée comme une application critique, nécessitant une disponibilité permanente, car elle dessert simultanément le service Finance, les ressources humaines et d'autres fonctions de gestion de la société. Pour y parvenir, la solution de recherche est déployée en grappes avec des options de basculement multi-sites ou sur un site unique.

Reprise sur sinistre et plan de continuité de l'activité :

En vue de respecter les principes de reprise sur sinistre et de continuité de l'activité, il est nécessaire de recourir à des systèmes matériels et logiciels redondants afin de réduire les risques de défaillance au maximum. Le principe général de cette conception consiste à disposer de composants en double permettant de créer une infrastructure redondante. Cela augmente la résilience en cas de défaillances de composants individuels et améliore la disponibilité du système pour les processus et les utilisateurs internes et externes.

ARCHITECTURE DE LA SOLUTION DE RECHERCHE

La fonction de recherche d'entreprise doit reposer sur une architecture robuste permettant d'atteindre les objectifs de la société. Elle se traduit par une architecture modulaire et multiniveaux, dans le respect des normes et principes énoncés ci-dessus. Elle doit en outre permettre de prendre en charge une solution de recherche couvrant les besoins actuels de l'organisation, être évolutive et pouvoir accélérer la mise sur le marché des produits et diminuer le coût total de possession.

Architecture conceptuelle

Le schéma°1 représente l'architecture conceptuelle d'une solution de recherche.

Une architecture conceptuelle est composée des éléments suivants :

Sources de contenu : Les sources de contenu comprennent l'intranet de la société, les sites Web, les serveurs de fichiers, les systèmes de gestion de contenus, tels que Documentum et FileNet, ainsi que les applications de l'entreprise qui couvrent tous les systèmes transactionnels, notamment les systèmes de gestion de base clients (CRM), de chaîne d'approvisionnement (SCM) et de planification des ressources de l'entreprise (ERP). En d'autres termes, elles englobent les différents systèmes d'une entreprise qui constituent les sources de données dans lesquelles les utilisateurs sont amenés à effectuer des recherches.

Services de connectivité : La couche des services de connectivité fournit les principaux adaptateurs permettant d'accéder aux sources de contenu sous-jacentes. Elle inclut deux types d'adaptateurs : les connecteurs et les accès fédérés. Les connecteurs permettent le fonctionnement de différents adaptateurs technologiques nécessaires à l'établissement d'une connexion à des sources de contenu. Les connecteurs d'accès fédérés permettent d'accéder à des contenus tiers.

Services de métadonnées : La couche des services de métadonnées fournit d'autres services à valeur ajoutée en matière de contenus. Ces services sont de deux types : mappage des métadonnées et dictionnaires de données. Les services de mappage des métadonnées permettent de définir les données provenant de systèmes divers. Ils permettent également le mappage individuel entre des éléments communs via la création d'un dictionnaire des données de l'entreprise couvrant plusieurs systèmes. Ce dictionnaire est composé d'un ensemble de définitions de métadonnées et de représentations d'éléments de données.

Accès fédérés : La couche des accès fédérés fournit une structure aux courtiers de requête qui permettent d'obtenir des informations provenant des systèmes transactionnels sous-jacents. L'accès aux données de l'entreprise requiert la création d'un index universel pour la sémantique.

Exemple d'architecture pour solution de recherche

Le schéma°2 présente un exemple d'architecture pour solution de recherche découlant de l'architecture conceptuelle décrite ci-dessus.

Les différentes couches de cette architecture sont les suivantes :

Interface utilisateur

Cette couche contrôle les interactions des utilisateurs via un navigateur Web ou même un appareil mobile. Elle détermine les écrans et les données associées qui sont présentés aux utilisateurs.

L'interface utilisateur est créée de façon dynamique en langage XML par la couche d'application. La présentation adéquate est délivrée par un fichier XSLT, en fonction du périphérique utilisé pour la requête. Présentation et contenu sont ainsi séparés et les responsabilités clairement réparties. Par ailleurs, cette couche est étroitement liée aux composants de sécurité (autorisation et authentification), assurant que les utilisateurs reçoivent uniquement les résultats de recherches pour lesquels ils possèdent une autorisation appropriée.

Les principaux éléments permettant d'optimiser l'utilisation d'une solution de recherche sont les suivants : une présentation basée sur les rôles, différents modes d'affichage, de modification et d'assistance contextuelle, ainsi que la mise en forme des résultats.

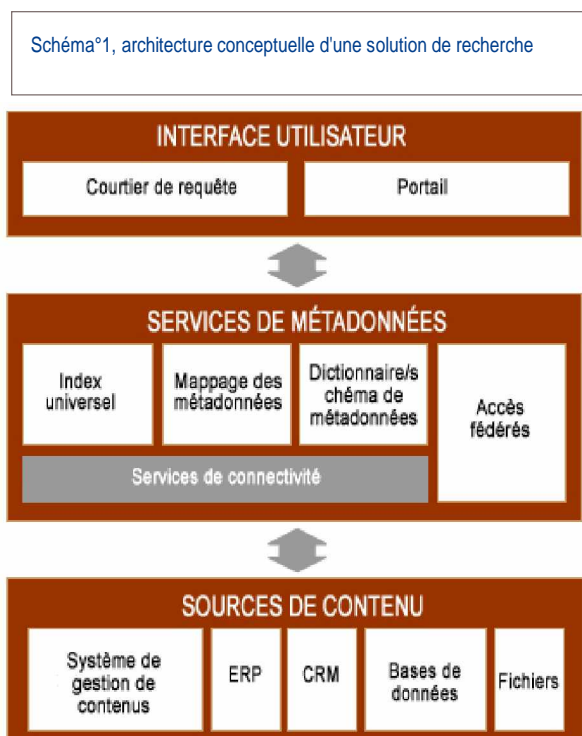
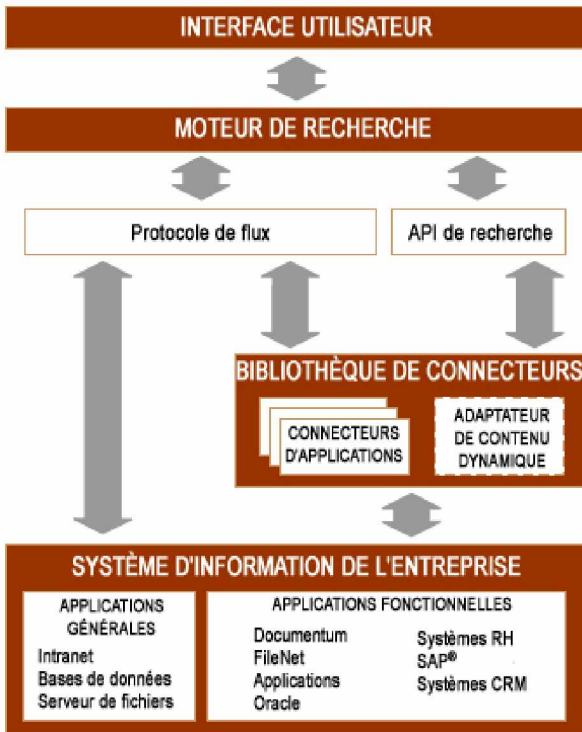


Schéma 2, exemple d'architecture pour une solution de recherche



La plupart des entreprises possèdent une foule de documents et de zones de stockage de contenus Web, disséminés sur leur intranet. Ces zones de stockage font appel à diverses technologies conçues pour l'hébergement du contenu. L'hébergement inclut les mécanismes de sécurité requis, les catalogues d'index individuels et le stockage du contenu en arrière-plan. L'architecture de la solution de recherche doit apporter des capacités dans trois domaines :

- **Prétraitement des requêtes :** Le préprocesseur de requêtes est un processus central pouvant être appelé à partir de divers emplacements. Sa fonction consiste à analyser la requête, à valider son orthographe, à la traduire dans d'autres langages, à identifier des synonymes, des concepts et le type de requête le mieux adapté à la zone de stockage de contenu associée ; par exemple, mot clé, concept, opérateurs booléens et texte complet.
- **Adaptateur de catalogue :** Les adaptateurs de catalogue permettent la mise en contact de la requête initiale avec les catalogues d'index de contenus en arrière-plan. Un index unique pour toutes les zones de stockage de contenus ne suffit pas à fournir les caractéristiques de sécurité nécessaires aux agents dédiés à la sécurité des informations.

À l'inverse, la méthode de mise en œuvre recommandée est d'avoir des adaptateurs capables de transmettre les requêtes au catalogue de contenu principal pour le compte de la personne ayant soumis la requête.

- **Agrégateur de résultats :** Ce domaine regroupe des résultats correspondant à la requête et issus de plusieurs catalogues afin de les présenter aux utilisateurs. Les pages de résultats doivent permettre de commenter chacun des résultats obtenus, de proposer une requête de recherche alternative et d'appliquer une classification des résultats par thème ou par catégorie au moteur de requête.

Moteur de recherche

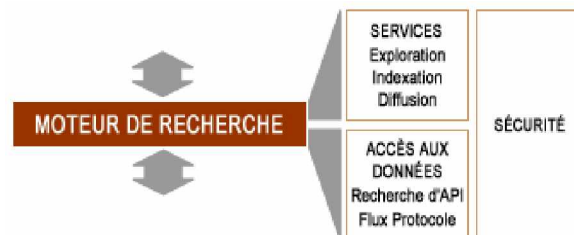
Le moteur de recherche est un produit matériel et/ou logiciel proposant une fonctionnalité de recherche à partir d'un champ unique pour l'ensemble de l'entreprise. Il doit être capable d'explorer des centaines de formats de fichiers, rédigés dans différents langages et peut indexer des millions de documents.

Il doit comprendre des fonctionnalités de sécurité afin d'assurer que les utilisateurs ne peuvent accéder qu'aux documents qu'ils sont autorisés à consulter.

Les fonctions essentielles d'un moteur de recherche englobent notamment (voir schéma 3) :

- **Distribution :** La fonction de distribution propose une interface de recherche standard pouvant être hébergée par le moteur de recherche par défaut. Cette interface doit être personnalisée si vous ne souhaitez pas utiliser la feuille de style XSLT. Les fonctionnalités de recherche supplémentaires, telles que les correspondances, les synonymes et les filtres, permettent de mettre en avant des pages Web spécifiques dans les résultats de recherche. Les synonymes suggèrent d'autres termes ou expressions correspondant aux requêtes de recherche. D'autres fonctionnalités classent les informations par catégories appelées "collections", ce qui permet de mieux répondre aux besoins spécifiques des utilisateurs en termes de recherche.

Schéma 3, fonctionnement du moteur de recherche



- **Exploration** : Les explorateurs sont des agents qui envoient des requêtes et récupèrent les documents des serveurs Web afin de les indexer automatiquement. L'exploration de données peut être effectuée à l'aide de deux méthodes : la découverte de données et la récupération de données. Il est possible d'explorer des sources de données diverses par le biais d'une interface de configuration. Tous les contenus sont regroupés dans un index principal qui est actualisé à chaque nouvelle exploration.

L'exploration est un processus qui se déroule en trois étapes : découverte initiale (nouvelles URL à explorer), indexation et exploration en continu. L'explorateur analyse tous les nouveaux documents afin de repérer les liens. L'exploration de ces liens peut être effectuée immédiatement ou planifiée à une date ultérieure. Cet explorateur peut traiter les contenus sécurisés et gère les communications S-HTTP (secure hypertext transfer protocol). L'explorateur du moteur de recherche est capable de traiter les authentifications de base, les authentifications NT LAN Manager (NTLM), les cookies personnalisés et les accès par formulaire. Le moteur explore également le contenu des bases de données, notamment Oracle, SQL Server, MySQL, IBM DB2 et Sybase. Les types de données auxquelles l'explorateur n'a pas accès peuvent être ajoutés directement dans le moteur de recherche en utilisant le format XML.

- **Indexation** : De nombreuses méthodes peuvent être appliquées pour indexer des données en amont :
 - Flux ajouté au système de recherche utilisant des adaptateurs standard ou personnalisés.
 - Accès direct au système de recherche (seuls les adaptateurs standard fournis avec le système peuvent accéder directement à ces données).
 - Flux de données HTTP ou non HTTP.
- **Flux** : Certains types de documents sont plus facilement intégrés au moteur de recherche à l'aide de flux qu'avec les liens figurant sur les pages Web explorées, notamment :
 - Les documents que l'explorateur ne parvient pas à trouver. Par exemple, les enregistrements d'une base de données ou les fichiers provenant d'un système qui ne prend pas en charge le format Web.
 - Les documents pouvant être explorés, mais qu'il est préférable d'explorer à un autre moment, en dehors des périodes planifiées dans le moteur de recherche.

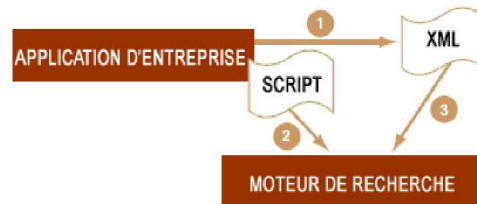
- Les documents pouvant être explorés mais qui ne disposent d'aucun lien sur le site Web qui permettrait à l'explorateur de les découvrir lors d'une nouvelle exploration.
- Les documents pouvant être explorés mais dont le transfert par le biais de flux est plus rapide en raison des dysfonctionnements du réseau ou du serveur Web.

Les types de flux à envisager sont les suivants :

- Les flux de contenus, composés d'URL des contenus associés. Ils peuvent également contenir des métadonnées, comme la date de dernière modification de l'enregistrement par exemple. Toutes les sources de données peuvent être définies comme des flux de contenus.
- Les flux du Web : dans ce cas, les enregistrements sont composés d'URL, non pas de contenus. Ces URL sont explorées de façon habituelle. Dans de tels flux, le terme "Web" est utilisé comme nom de la source de données. Toutes les URL doivent contenir un nom de domaine complet (FQDN) dans la section d'hébergement de l'URL.
- Les flux de bases de données : les connecteurs personnalisés peuvent être configurés pour extraire les enregistrements d'une base de données afin de créer un index.

Le processus de configuration des flux de moteurs de recherche est illustré par le schéma 4.

Schéma 4, flux d'API pour la recherche de contenus



- 1 Le script exporte les données au format XML afin de générer un fichier de flux.
- 2 Il envoie le fichier XML au moteur de recherche.
- 3 Le moteur de recherche importe le fichier XML et indexe son contenu.

• **Sécurité du moteur de recherche** : Le moteur de recherche doit renforcer le système de sécurité de l'entreprise. Tous les services doivent être exécutés derrière un pare-feu, limitant le nombre de ports qui doivent être ouverts pour communiquer avec le moteur de recherche à travers le pare-feu. Le moteur de recherche explore et indexe à la fois les documents publics et confidentiels. Un package de sécurité peut également être mis en place pour étendre le dispositif au niveau des documents.

Pour délivrer des résultats de recherche appropriés à tous les utilisateurs, il est nécessaire d'appliquer des filtres aux résultats obtenus dans l'index de recherche. Les contenus sont répartis en deux catégories : contenu public et contenu intégral. Les utilisateurs peuvent rechercher dans les contenus publics uniquement ou dans le contenu public et dans les documents sécurisés de l'index, en fonction des spécifications établies par l'administrateur de l'index. Par défaut, tous les contenus sont sécurisés, afin de s'assurer que les utilisateurs ne puissent pas visualiser des contenus qu'ils n'ont pas le droit de consulter.

Le moteur de recherche doit prendre en charge les authentifications de base et NTLM (NT LAN Manager) ainsi que les authentifications par formulaire. Ces protocoles d'authentification sont traités séparément par le moteur de recherche et requièrent une configuration spécifique que l'administrateur doit définir. Chacune de ces méthodes permet d'explorer et d'indexer les documents protégés des sites intranet et de réaliser l'authentification des utilisateurs effectuant une recherche en rapport avec ces documents.

Une API d'autorisation de recherche fait communiquer l'API d'autorisation du moteur de recherche et le serveur de l'entreprise qui fournit des services de contrôle d'accès, également dénommés connecteurs d'accès (AC). Les connecteurs d'accès constituent une couche entre le moteur de recherche et le système de contrôle d'accès de la société.

Bibliothèque de connecteurs

Cette couche logicielle constitue la couche de messagerie et/ou d'accès à différentes sources de données de l'entreprise, allant des sources de données structurées aux applications de la société. En règle générale, une plateforme d'intégration ou des connecteurs ou adaptateurs personnalisés sont conçus pour intégrer les sources de données.

La création de connecteurs personnalisés peut également être effectuée de la manière suivante :

- Conversion des données au format XML spécifié dans le flux du moteur de recherche.

- Transfert des données XML au moteur par le biais du protocole HTTP.

Cette couche forme la base sous-jacente permettant de diffuser l'information dans une entreprise, quels que soient la source et le format de stockage. Cette couche permet en effet le couplage des services avec les adaptateurs d'intégration provenant de sources de données diverses.

Les métadonnées définies dans cette couche peuvent également fournir l'infrastructure nécessaire à la consolidation des informations provenant de sources différentes dans l'entreprise. Cela permet également de découvrir et de renvoyer rapidement des éléments ayant une correspondance sémantique. Ce métamodèle unifié établit ainsi des définitions de métadonnées cohérentes et contextuellement correctes, permettant la consolidation d'informations disparates. Il améliore également les résultats de recherche en fournissant des informations, notamment des liens déterminés en fonction du contexte pointant vers des données structurées et des applications d'entreprise.

Sources et types d'informations

De nos jours, les informations sont disséminées dans toute l'entreprise. Elles sont disponibles dans des formats structurés ou non structurés, dans des référentiels protégés ou non (voir Tableau 1). Les moteurs de recherche sont capables d'explorer les informations et de les indexer afin de faciliter leur accès, améliorant ainsi les résultats de recherche. Il est donc impératif de classer les sources de données de la même manière et d'utiliser les mécanismes de renvoi appropriés.

Tableau 1. Sources et types d'informations

	PROTÉGÉ	NON PROTÉGÉ
Structuré	<p>Systèmes de gestion de base de données (Oracle, SQL Server, Sybase, Informix)</p> <p>Systèmes de gestion de documents (Documentum, FileNet, etc.)</p> <p>Applications commerciales de type ERP et CRM</p>	<p>Systèmes de gestion de base de données (mySQL)</p>
Non structuré	<p>Systèmes de fichiers réseau/systèmes de fichiers</p> <p>Documents Microsoft Office</p> <p>Fichiers image, vidéo et audio</p>	<p>Documents Open office</p> <p>Systèmes de fichiers/fichiers texte du serveur de fichiers</p> <p>E-mail</p>

Informations structurées et non structurées : Ces deux types d'informations se différencient principalement par l'utilisation d'un modèle d'information prédéfini. Les informations structurées sont typiquement des éléments que l'on peut associer à un format prédéfini, tel que le format HTML, ou pouvant être insérés dans un tableau, dans le cas des bases de données. Il est possible également de les considérer en termes de métadonnées : des informations sur les informations. À l'inverse, les données non structurées n'ont pas de format prédéfini et existent ainsi sous forme libre. Citons notamment à titre d'exemple, les fichiers texte et les documents PDF. Dans un environnement non structuré, la solution de recherche doit être capable d'extraire l'information.

Référentiels protégés et non protégés : De nos jours, les informations des entreprises figurent dans des systèmes ERP, CRM et de gestion de connaissances, ainsi que sur des portails statiques et dynamiques. Dans le cadre de la recherche d'informations, certains de ces systèmes fournissent des formats protégés pour la publication des informations. D'autres emploient des formats non protégés, tels que les données XML, qui font d'ailleurs de plus en plus d'adeptes. Par ailleurs, si le texte reste dominant dans de nombreuses applications, la pertinence des autres types de supports est croissante, notamment les images, le son et la vidéo. Il est donc nécessaire de combiner efficacement l'extraction automatique de texte, de métadonnées (créées manuellement, en général) et l'extraction de données multimédia basée sur le contenu.

SOLUTION DE RECHERCHE, UNE OPPORTUNITÉ À SAISIR

Adopter une solution de recherche dans une entreprise peut favoriser la productivité des employés et permettre de prendre des décisions en connaissance de cause. Cependant, l'intégration de la solution de recherche dans l'environnement de l'entreprise suppose une compréhension précise des capacités et des exigences en termes d'architecture. Une architecture soigneusement planifiée et mise en œuvre est la clé permettant de bénéficier pleinement des avantages considérables que procure une solution de recherche.

Pour en savoir plus sur la manière dont votre entreprise peut tirer parti de nos solutions, [contactez-nous](#).

CONSEIL EN GESTION GLOBALE ET EN TECHNOLOGIE POUR LES ENTREPRISES D'AUJOURD'HUI

BearingPoint est une société de conseil, leader en gestion globale et en technologie, qui compte parmi ses clients, Global 2000 et de nombreuses institutions de service public parmi les plus grandes du monde. Nos experts aident des organisations dans le monde entier à définir une ligne de conduite permettant d'atteindre leurs objectifs et de créer une valeur commerciale. En alignant leurs processus métier et leurs systèmes d'information, nous aidons nos clients à se positionner en leader face à la concurrence, grâce à des résultats rapides. Pour en savoir plus, contactez-nous au 1.866.661.FIND (+1.603.589.4089 en dehors des États-Unis et du Canada) ou visitez notre site Web à l'adresse www.bearingpoint.com.

