

CRÉATION ET GESTION D'UN INDEX DANS UN DOCUMENT PDF

Vous pouvez réduire le temps nécessaire à effectuer les recherches dans un long document PDF en incorporant dans celui-ci un index des mots. Acrobat met beaucoup moins de temps à rechercher des mots dans un index que dans un document complet. L'index incorporé est fourni dans les exemplaires diffusés ou partagés du document PDF. Les utilisateurs procèdent de la même façon pour effectuer des recherches dans des documents PDF dotés d'un index incorporé ou non ; aucune étape supplémentaire n'est requise.

Ajout d'un index à un document PDF

1. Le document étant ouvert dans Acrobat, choisissez Options avancées > Traitement du document > Gérer l'index incorporé.
2. Dans la boîte de dialogue Gérer l'index incorporé, cliquez sur Incorporer l'index.
3. Lisez les messages qui s'affichent, puis cliquez sur OK.

Remarque : *Dans Outlook et Lotus Notes, vous avez la possibilité d'incorporer un index au moment de la conversion des messages électroniques ou des dossiers de courrier au format PDF. Cette option est particulièrement recommandée dans le cas de dossiers contenant de nombreux messages électroniques.*

Mise à jour ou suppression de l'index incorporé dans un document PDF

1. Choisissez Options avancées > Traitement du document > Gérer l'index incorporé.
2. Cliquez sur Mettre à jour l'index ou Supprimer l'index.
3. Cliquez sur OK dans le message de confirmation.

A PROPOS DE LA FONCTION CATALOG

Vous pouvez définir un groupe spécifique de documents PDF sous forme de **catalogue** et créer un index unifié pour la série complète de documents. Lorsqu'un utilisateur recherche des informations spécifiques dans les fichiers PDF catalogués, l'index accélère considérablement l'opération.

Lorsque vous diffusez ou publiez la série de documents sur un CD ou un site Web, vous pouvez ajouter l'index aux fichiers PDF.

Vous pouvez cataloguer des documents écrits en caractères romains, chinois, japonais ou coréens. Les éléments pouvant être catalogués incluent le texte, les commentaires, les signets, les champs de formulaire et les balises d'un document, les métadonnées de document et d'objet, les pièces jointes, les informations sur le document, les signatures numériques, les métadonnées d'image au format XIF et les propriétés personnalisées du document.

PRÉPARATION D'UN DOCUMENT PDF POUR L'INDEXATION

Commencez par créer un dossier contenant les fichiers PDF à indexer. Tous les fichiers PDF doivent être complets au niveau du contenu comme des fonctions électroniques telles que les liens, les signets et les champs de formulaire. Si les fichiers à indexer comprennent des documents numérisés, assurez-vous que le texte est disponible pour la recherche. Fractionnez les documents volumineux en fichiers plus petits, par chapitres, pour optimiser les performances de recherche. Vous pouvez également ajouter des informations aux propriétés du document d'un fichier afin d'améliorer les possibilités de recherche dans le fichier.

Avant d'indexer une série de documents, il est essentiel que vous configuriez la structure des documents sur le lecteur de disque ou sur le volume du serveur réseau et que vous vérifiiez les noms de fichiers multiplates-formes. Les noms de fichiers risquent d'être tronqués et difficilement récupérables lors d'une recherche multiplate-forme. Pour éviter ce problème, tenez compte des recommandations suivantes :

- Renommez les fichiers, les dossiers et les index conformément à la convention de dénomination de fichiers MS-DOS (huit caractères ou moins suivis d'une extension de trois caractères) et ce, tout particulièrement si vous projetez de diffuser la série de documents et l'index sur un disque CD-ROM formaté ISO 9660.

- Supprimez des noms de fichiers et de dossiers les caractères étendus, tels que les caractères accentués et non anglais. (La police utilisée par la fonction Catalog ne prend pas en charge les codes de caractères compris entre 133 et 159.)
- Abstenez-vous d'imbriquer trop profondément les dossiers ou d'utiliser des noms de chemins dépassant 256 caractères pour les index qui seront utilisés sous Mac OS.
- Si vous utilisez Mac OS avec un serveur de réseau local (LAN) OS/2, configurez LAN Server Macintosh (LSM) de manière à appliquer les conventions de dénomination de fichiers MS-DOS ou indexez seulement les volumes FAT. (Les volumes HPFS peuvent contenir des noms de fichiers longs irrécupérables.)

Si la structure du document comprend des sous-dossiers à ne pas indexer, vous pouvez exclure ces derniers au cours de l'indexation.

AJOUT DE MÉTADONNÉES AUX PROPRIÉTÉS DU DOCUMENT

Pour faciliter la recherche dans un fichier PDF, vous pouvez insérer des informations sur le fichier dans les propriétés du document (informations appelées **métadonnées**). (Pour afficher les propriétés du fichier PDF actif, choisissez Fichier > Propriétés, puis cliquez sur l'onglet Description.)

(Windows) Vous avez également la possibilité de saisir et de lire ces informations depuis le bureau. Cliquez avec le bouton droit de la souris sur le document dans l'Explorateur Windows, choisissez Propriétés et cliquez sur l'onglet PDF. Toutes les informations que vous saisissez ou que vous modifiez dans cette boîte de dialogue figurent également dans le panneau Description de la boîte de dialogue Propriétés du document lorsque vous ouvrez le fichier.

Lorsque vous complétez les propriétés du document, tenez compte des recommandations suivantes :

- Choisissez un titre représentatif. Le nom de fichier du document devrait figurer dans la boîte de dialogue des résultats de la recherche.
- Utilisez toujours la même option (champ) pour les informations similaires. Par exemple, ne choisissez pas le champ Sujet afin d'ajouter un terme important pour certains documents et le champ Mots-clés pour d'autres.
- Utilisez systématiquement le même terme pour désigner une information donnée. Par exemple, n'utilisez pas **biologie** pour certains documents et **sciences de la vie** pour d'autres.
- Utilisez le champ Auteur pour identifier le groupe responsable du document. Par exemple, l'auteur d'un document relatif au recrutement du personnel pourrait être le service des ressources humaines.
- Si vous identifiez vos documents par des numéros de référence, ajoutez-les dans l'option Mots-clés. Par exemple, l'ajout de **doc#=m234** à l'option Mots-clés référence un document spécifique dans une série de plusieurs centaines de documents relatifs à un sujet particulier.
- Pour classer les documents par type, utilisez l'option Sujet ou Mots-clés. Par exemple, indiquez **rapport** comme sujet et **mensuel** ou **hebdomadaire** comme mot-clé d'un document.

Si vous êtes expérimenté dans l'utilisation du format Adobe PDF, vous pouvez définir des champs de données personnalisés (tels que Type de document, Numéro du document et Identifiant du document) lors de la création de l'index. Cette méthode s'adresse uniquement aux utilisateurs chevronnés et n'est pas traitée dans l'aide complète d'Acrobat.

CRÉATION D'UN INDEX POUR UNE SÉRIE DE DOCUMENTS

Lorsque vous constituez un nouvel index, Acrobat crée un fichier doté de l'extension .pdx et un nouveau dossier de support, lequel contient un ou plusieurs fichiers dotés de l'extension .idx. Les fichiers IDX contiennent les entrées d'index. Ces fichiers doivent tous être accessibles pour les utilisateurs souhaitant parcourir l'index.

Boîtes de dialogue relatives à la création d'index

1. Choisissez Options avancées > Traitement du document > Index de texte intégral avec Catalog, puis cliquez sur Nouvel index.
2. Dans le champ Titre, indiquez le nom du fichier index.
3. Dans la zone Description, saisissez quelques mots concernant le type ou la fonction de l'index.
4. Choisissez Options, sélectionnez les options avancées à appliquer à l'index, puis cliquez sur OK.
5. Sous Dossiers à inclure, cliquez sur Ajouter, sélectionnez un dossier contenant une partie ou la totalité des fichiers PDF à indexer, puis choisissez OK. Pour ajouter d'autres dossiers, recommencez cette étape.

Remarque : *Tout dossier imbriqué sous un dossier inclus dans l'opération d'indexation est également compris dans la procédure. Vous pouvez ajouter des dossiers situés sur différents serveurs ou lecteurs de disque, à condition que vous ne projetiez pas de déplacer l'index ou tout élément contenu dans la série de documents.*

6. Sous Dossiers à exclure, cliquez sur Ajouter, puis sélectionnez tout dossier imbriqué contenant des fichiers PDF à ne pas indexer. Cliquez sur OK et, le cas échéant, recommencez cette étape.
7. Vérifiez les sélections effectuées. Pour apporter des changements aux listes de dossiers à inclure ou à exclure, sélectionnez le dossier à modifier et cliquez sur Supprimer.
8. Cliquez sur Constituer, puis spécifiez l'emplacement du fichier index. Cliquez sur Enregistrer, puis :
 - Cliquez sur Fermer une fois l'indexation terminée.
 - Cliquez sur Arrêter pour annuler la procédure d'indexation.

Remarque : *Si vous arrêtez l'opération, vous ne pouvez pas reprendre la même session d'indexation, mais vous n'avez pas besoin de recommencer à zéro. En effet, les options et les sélections restent intactes. Vous pouvez cliquer sur Ouvrir un index, sélectionnez l'index partiellement créé et en vérifier le contenu.*

Si les noms de chemins longs sont tronqués dans les zones Dossiers à inclure et Dossiers à exclure, maintenez le pointeur enfoncé sur les points de suspension (...) jusqu'à ce qu'une info-bulle affiche le chemin d'accès complet.

Boîte de dialogue des options d'indexation

Exclure les nombres : Cochez cette case si vous souhaitez exclure de l'index tous les nombres figurant dans le texte du document. L'exclusion de nombres peut réduire considérablement la taille d'un index, accélérant les recherches.

Ajouter un identifiant aux fichiers Adobe PDF créés dans Acrobat 1.0 Cochez cette option si la série de documents comprend des fichiers PDF créés avant Acrobat 2.0, lesquels n'ajoutent pas automatiquement des numéros d'identification. Ces identifiants sont nécessaires lorsque les noms de fichiers Mac OS longs sont raccourcis lors de leur conversion en noms de fichiers DOS. Les versions 2.0 et ultérieures d'Acrobat ajoutent automatiquement des identifiants.

Ne pas signaler les documents modifiés lors des recherches : Lorsque cette option n'est pas activée, un message s'affiche pendant la recherche dans les documents modifiés depuis la dernière constitution de l'index.

Propriétés personnalisées : Utilisez cette option pour inclure les propriétés de document personnalisées dans l'index. Seules les propriétés de document personnalisées qui se trouvent déjà dans les documents PDF inclus dans l'index sont indexées. Indiquez la propriété, sélectionnez un élément dans le menu Type, puis cliquez sur Ajouter. Ces propriétés s'affichent comme option de recherche dans les menus déroulants

de critères supplémentaires de la fenêtre Recherche avancée lorsque vous effectuez une recherche dans l'index résultant. Si, par exemple, vous saisissez la propriété personnalisée Nom du document et choisissez la propriété Chaîne dans le menu Type, un utilisateur parcourant l'index peut ensuite effectuer la recherche dans la propriété personnalisée en sélectionnant Nom du document dans le menu Critères supplémentaires.

Remarque : *Lorsque vous créez des champs personnalisés dans une application Microsoft Office dont l'option Convertir les informations sur le document est activée dans l'application PDFMaker, les champs sont conservés dans tout fichier Adobe PDF que vous créez par la suite.*

Champs XMP Utilisez cette option pour inclure des champs XMP personnalisés. Les champs XMP personnalisés sont indexés et s'affichent dans les menus de critères supplémentaires. Ils font partie des éléments dans lesquels la recherche s'effectue.

Mots éliminés : Permet d'exclure des mots spécifiques (500 au maximum) des résultats de la recherche dans l'index. Saisissez le mot, cliquez sur Ajouter et recommencez l'étape le cas échéant. L'exclusion de mots de l'index permet de réduire la taille de l'index, de l'ordre de 10 à 15 %. Les mots éliminés de l'indexation ne doivent pas dépasser 128 caractères et respectent la casse. Certains mots courants, tels que le, et et à sont déjà exclus de la recherche dans les index, il est donc inutile de les rajouter.

Remarque : *Pour empêcher les utilisateurs de tenter de rechercher des expressions contenant ces mots, dressez la liste des mots qui ne sont pas indexés dans le fichier Lisezmoi ou Ouvrez-moi de Catalog.*

Balises structurales Ce bouton permet de rendre disponibles pour la recherche des noeuds de balises élément contenus dans des documents dotés d'une structure logique balisée.

Remarque : *Les options Propriétés personnalisées, Mots éliminés et Balises s'appliquent uniquement à l'index actif. Pour appliquer ces paramètres de manière globale à tout index créé, modifiez les paramètres par défaut des champs personnalisés, des mots éliminés de l'indexation et des balises dans le panneau Catalog de la boîte de dialogue Préférences.*

FICHIERS LISEZMOI OU OUVREZ-MOI DE CATALOG

Il est souvent judicieux de créer un fichier Lisezmoi ou Ouvrez-moi distinct que vous incluez dans le dossier avec l'index. Ce fichier peut fournir aux lecteurs des détails sur votre index, par exemple :

- Type de documents indexés
- Options de recherche prises en charge
- Personne à contacter ou un numéro de téléphone à appeler en cas de questions
- Liste des nombres ou termes exclus de l'indexation
- Liste des dossiers contenant les documents inclus dans un index de réseau local ou la liste des documents inclus dans un index pour disque. Vous pouvez également inclure une brève description du contenu de chaque dossier ou document.
- Liste des valeurs associées à chaque document si vous attribuez des valeurs de champs dans les informations sur le document.

Si un catalogue contient un nombre particulièrement important de documents, envisagez d'inclure un tableau indiquant les valeurs attribuées à chaque document. Ce tableau peut faire partie de votre fichier Lisezmoi ou Ouvrez-moi, ou faire l'objet d'un document distinct. Lors de la mise en place de l'index, consultez ce tableau afin de conserver une certaine cohérence entre les documents.

VÉRIFICATION D'UN INDEX

Vous pouvez mettre à jour, reconstituer ou vider un index existant.

1. Choisissez Options avancées > Traitement du document > Index de texte intégral avec Catalog, puis cliquez sur Ouvrir un index.
2. Localisez et sélectionnez le fichier de définition d'index (fichier PDX) correspondant à l'index, puis cliquez sur Ouvrir.
3. Si l'index a été créé dans Acrobat 5.0 ou version antérieure, cliquez sur Créer une copie pour définir un nouvel index (sans écraser l'ancienne version) ou sur Ecraser l'ancien index pour écraser l'index précédent.
4. Dans la boîte de dialogue Définition de l'index, apportez les modifications voulues, puis sélectionnez la fonction qu'Acrobat doit effectuer :

Constituer... Crée un nouveau fichier IDX contenant les informations existantes et les met à jour en ajoutant des entrées et en signalant les entrées modifiées ou obsolètes comme incorrectes. Si vous apportez de nombreuses modifications ou utilisez plusieurs fois cette option au lieu de créer un nouvel index, la durée des opérations de recherche peut augmenter.

Reconstituer Crée un nouvel index, écrasant le dossier d'index existant et son contenu (les fichiers IDX).

Vider Supprime le contenu de l'index (les fichiers IDX) sans supprimer le fichier index lui-même (PDX).

CATALOG, PRÉFÉRENCES

Vous pouvez configurer les préférences d'indexation s'appliquant de manière globale à tous vos futurs index. Vous avez la possibilité d'annuler ces préférences pour un index particulier en sélectionnant d'autres options lors de la constitution de l'index.

Pour ouvrir les préférences de Catalog, choisissez Edition > Préférences (Windows) ou Acrobat > Préférences (Mac OS), puis sélectionnez Catalog dans la liste Catégories. La plupart des options sont identiques à celles décrites dans la procédure de constitution d'index.

L'option Appliquer la norme ISO 9660 aux dossiers s'avère utile pour éviter de modifier des noms de fichiers PDF longs en noms de fichiers MS-DOS lors de la préparation des documents pour l'indexation. Cependant, vous devez toujours appliquer les conventions de dénomination de fichiers MS-DOS aux noms de dossiers (8 caractères au maximum), même si cela est inutile pour les noms de fichiers.

MISES A JOUR D'INDEX PLANIFIEES

Il est possible de planifier le moment et la fréquence de la constitution, la reconstitution, la mise à jour et le vidage automatiques d'un index à l'aide de la fonction Catalog et d'un fichier PDX de lot Catalog (.bpdX). Un fichier BPDX est un fichier texte contenant une liste de chemins d'accès aux fichiers index Catalog en fonction de la plate-forme et leurs indicateurs. Vous pouvez utiliser une application de planification (Planificateur de tâches pour Windows ou Scheduler pour Mac OS) afin d'afficher le fichier BPDX dans Acrobat. Acrobat recrée alors l'index en fonction des indicateurs contenus dans le fichier BPDX.

Pour plus de détails sur la planification d'une mise à jour d'index, recherchez BPDX sur la page Support du site Web d'Adobe.

DÉPLACEMENT DE SÉRIES DE DOCUMENTS ET DES INDEX ASSOCIÉS

Vous pouvez ouvrir et tester une série de documents indexés sur un lecteur de disque local, puis la placer sur un serveur réseau ou sur un disque. Une définition d'index contient des chemins relatifs entre le fichier de définition de l'index (PDX) et les dossiers contenant les documents indexés. Si ces chemins relatifs restent inchangés, il est inutile de reconstituer l'index après le déplacement de la série de documents indexée. Si le fichier PDX et les dossiers contenant les documents indexés se trouvent dans le même dossier, il vous suffit de déplacer le dossier pour conserver le même chemin relatif.

Si le chemin d'accès relatif change, vous devez créer un nouvel index après avoir déplacé la série de documents indexés. Vous pouvez toutefois continuer à utiliser le fichier PDF d'origine. Pour cela, vous devez d'abord déplacer les documents indexés, puis copier le fichier PDX dans le dossier dans lequel vous voulez créer l'index. Le cas échéant, modifiez les listes de répertoires et de sous-répertoires dans les zones Dossiers à inclure et Dossiers à exclure.

Si l'index ne se trouve pas sur le même lecteur ou sur le même volume de serveur que n'importe lequel des documents de la série à laquelle il est associé, vous rompez l'index en déplaçant les documents ou l'index. Si vous projetez de déplacer une série de documents vers un autre lecteur du serveur ou sur un CD-ROM, créez et constituez l'index à l'emplacement où se trouvent les documents.