

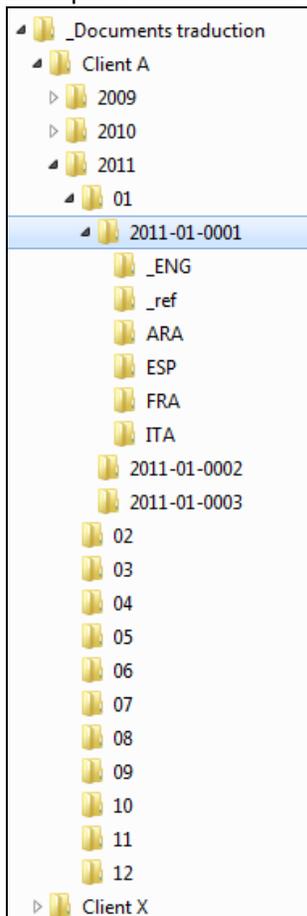
Bonnes pratiques d'archivage de documents multilingues

L'archivage de documents peut sembler facile à première vue, mais beaucoup de personnes échouent lamentablement à cette tâche. Un bon système d'archivage constitue la base de tout environnement de travail car votre efficacité en dépend directement et ce, peu importe les logiciels que vous utilisez. Le présent document expose des bonnes pratiques qui ont fait leurs preuves. En aucun cas nous ne prétendons que le présent document est la panacée dans le domaine; il n'est plutôt qu'un simple cumul de nombreuses années d'expérience dont nous avons bien voulu vous faire part.

Création d'une arborescence logique

L'arborescence représente l'épine dorsale de votre archivage. Avant de nommer vos fichiers, vous devez les classer correctement. Tout document est lié à un client, à une catégorie, à un domaine ou à tout autre regroupement qui s'applique à votre situation. Ce lien ou ce regroupement est représenté par des répertoires (dossiers).

Exemple d'arborescence structurée :



Le répertoire *_Documents traduction* représente la racine de l'arborescence et est précédé d'un sous-tiret. En informatique, le sous-tiret est un des caractères ayant préséance sur les lettres et les chiffres dans un classement alphabétique. En mettant un sous-tiret devant le répertoire-racine, ce dernier sera nécessairement classé en début de liste, ce qui facilite grandement l'accès aux fichiers. Pour les répertoires numérotés, assurez-vous d'avoir le même nombre de chiffres pour tous les répertoires. Dans l'exemple ci-dessus, les mois de l'année sont écrits avec deux chiffres (01, 02, 03, etc.). Si vous pensez créer entre 1 et 999 répertoires numérotés, vous devez adopter un système à trois chiffres (001, 002, etc.) et non pas 1, 2, 3, etc., car l'ordre logique des répertoires ne sera pas respecté.

- Sous le répertoire-racine, chaque client a son propre répertoire (Client A, Client X, etc.).
- Sous chaque répertoire-client, nous avons ajouté des répertoires-années. Chaque répertoire-année est précédé d'un chiffre afin que les mois apparaissent dans le bon ordre.
- Sous chaque répertoire-année, nous avons ajouté des répertoires-mois représentés par des chiffres.
- Sous chaque répertoire-mois, nous avons ajouté le numéro de la demande de traduction. Le numéro de demande devrait fournir de l'information pertinente et logique comme l'année, le mois ou autre élément qui vous semble pertinent.
- Sous chaque numéro de demande de traduction, nous avons ajouté les répertoires-langues ainsi que le répertoire « *_ref* » pour y verser tout document de référence correspondant à cette demande de traduction.

Ce qui nous intéresse plus particulièrement dans ce guide est la structure des répertoires-langues et les fichiers qui y résident.

Reprenons l'exemple de l'arborescence en page 1. Pour la demande 2011-01-0001, le répertoire-langue « *_ENG* » est précédé d'un sous-tiret, car il s'agit de la langue pivot (langue source), comme l'est souvent l'anglais. De cette façon, les traducteurs savent dans quel répertoire se trouvent les textes à traduire. Dans un contexte multilingue, et sans méthode de nommage rigoureuse, il peut être difficile de savoir quelle est la langue pivot.

Nommage de documents

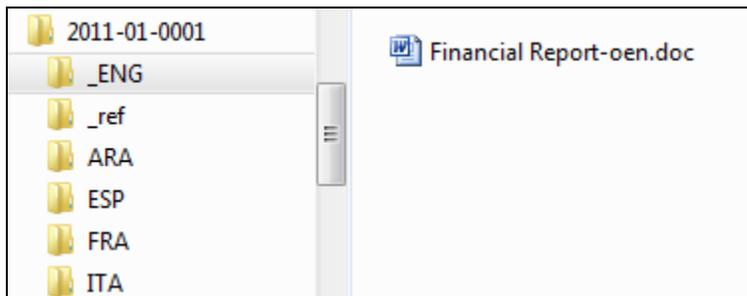
Voici quelques règles de base à respecter pour un classement optimal de vos documents :

- Les documents source et cible doivent toujours porter le même nom, ou à peu de différences près. Ne traduisez jamais le nom de vos documents;
- Les documents doivent porter un marqueur de langue à la fin plutôt qu'au début du nom pour éviter un tri alphabétique erroné. Le marqueur de langue devrait être séparé du nom du fichier par un sous-tiret ou un tiret. Évitez l'espace comme séparateur;
- Le nom des documents devrait être le plus court possible, représentatif et indiquer un maximum de détails sur le contenu (si vous devez ouvrir le document pour savoir ce qu'il contient, le nom de votre document n'est pas clair et vous perdez votre temps).

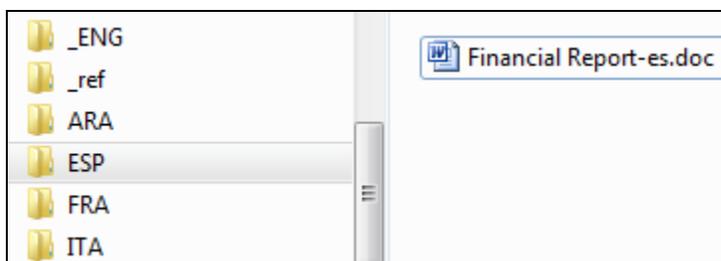
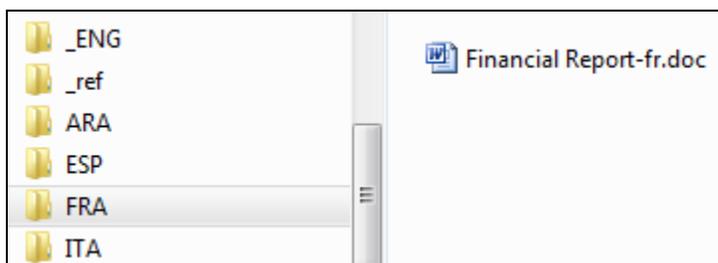
- Évitez tout caractère accentué dans le nom de vos fichiers et de vos répertoires. Les logiciels de compression (Winzip, 7zip, etc.) peuvent donner des erreurs à la compression ou à la décompression de répertoires ou de fichiers zippés.

Marqueurs de langues

Pour diverses raisons, il peut être utile de savoir quel est le document de départ par rapport aux documents cibles (traductions) en ne regardant que le nom des documents. Les marqueurs de langue sont donc essentiels. Voici un exemple :



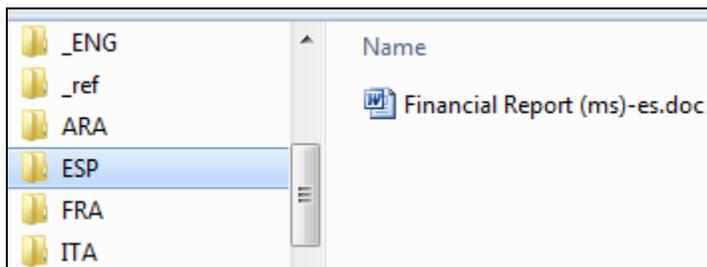
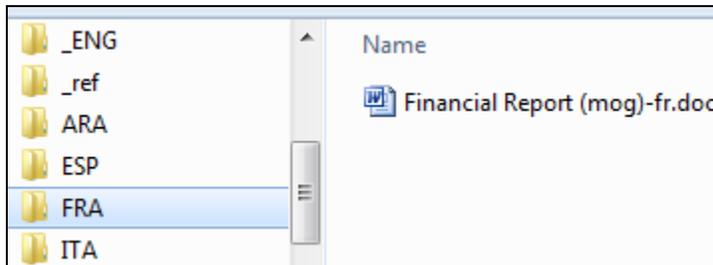
Dans cet exemple, la langue source est l'anglais, nous avons donc ajouté « oen » comme marqueur de langue, ce qui signifie « original English ». Les traductions qui découlent de ce document ne porteront évidemment pas le marqueur « o », mais uniquement leur marqueur de langue normal. Voici des exemples de documents traduits :



Le choix de marqueurs de langues est tout à fait arbitraire. Nous avons choisi des marqueurs composés de deux lettres pour éviter les noms de documents trop longs. Dans cet exemple, nos marqueurs sont français=fr, arabe=ar, espagnol=es, italien=it, anglais=en. Une copie du document source est copiée dans chaque répertoire et renommée selon la langue cible. Les traducteurs traduisent donc dans leur répertoire respectif.

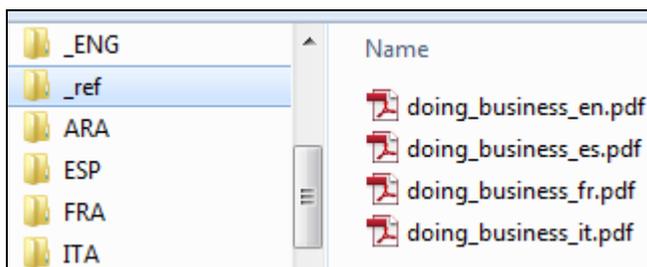
Initiales dans le nom de document

Les lettres entre parenthèses représentent les initiales des traducteurs. Encore une fois, il est recommandé de ne jamais placer ces informations au début du nom du document pour éviter que les documents soient séparés dans une liste alphabétique. L'ajout des initiales du traducteur directement dans le nom du document peut également donner des informations supplémentaires.



Documents de référence

Pour certaines demandes de traduction, vous pouvez disposer de documents de référence de tout type (lexiques, notes, documents, etc.). Dans cet exemple, le client a fourni quatre documents pour aider à la traduction. Nous avons renommé les documents de façon à ce qu'ils respectent notre nomenclature de documents.



Pour optimiser l'utilisation de ces références, nous vous recommandons de créer des bitextes à partir de ces documents (voir paragraphe suivant).

Création de bitextes à partir de documents multilingues

Une fois la traduction terminée, vous devez alimenter votre mémoire de traduction. Selon le produit que vous utilisez, vous devrez peut-être aligner les documents avant de les importer dans votre mémoire. Si tel est le cas, voici une façon de faire. Nous avons choisi de vous présenter l'outil AlignFactory pour générer les alignements qui serviront à aligner une mémoire de traduction multilingue.

AlignFactory permet d'automatiser l'appariement et l'alignement des documents dans des répertoires entiers, ce que ne permet pas AlignFactory Light (appariement manuel seulement). Il faut bien distinguer l'appariement de l'alignement proprement dit. Les exemples exposés ci-dessous sont basés sur l'automatisation de l'appariement et non pas sur l'appariement manuel. Le processus d'appariement est basé sur les noms de fichiers, d'où l'importance d'adopter une nomenclature logique et rigoureuse. Avant de se servir d'AlignFactory, vous devez créer des paires de langues avec lesquelles vous travaillez. Vous ne pouvez créer deux paires avec les mêmes langues, par exemple, Anglais-Français et Français-Anglais. Cela serait inutile car un bitexte est utilisé de façon bidirectionnelle. Pour les exemples suivants, nous avons créé quatre paires de langues en indiquant l'anglais comme langue source. Cela signifie donc que les langues indiquées dans le nom du bitexte commenceront toujours par le code ENG suivi du code de l'autre langue, par exemple, ENG-FRA_BT.xml.

Ajout des paires de langues dans AlignFactory.

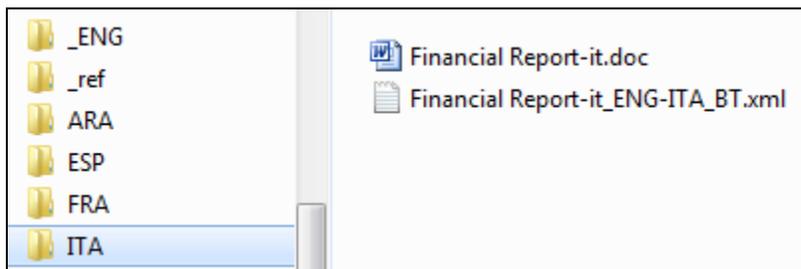
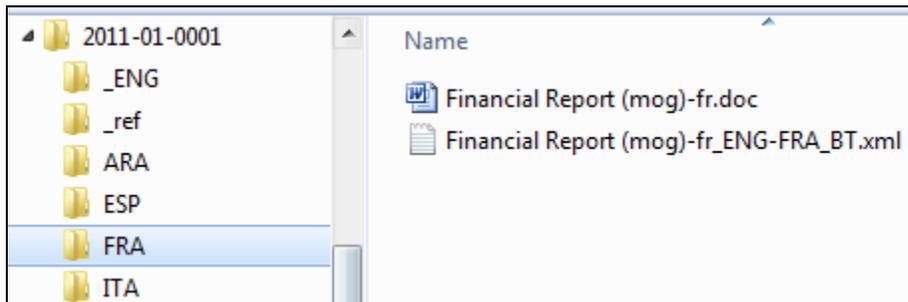
Langue de départ :		Langue d'arrivée :	
Anglais	<Générique>	Anglais	<Générique>
Langue de départ	Langue d'arrivée		
Anglais (Générique)	Français (Générique)		
Anglais (Générique)	Arabe (Générique)		
Anglais (Générique)	Espagnol (Générique)		
Anglais (Générique)	Italien (Générique)		

Une fois les paires de langues paramétrées, il est inutile de les modifier même dans le cas où l'anglais ne serait pas la langue source. L'utilisation de marqueurs précis dans le nom de documents revêt donc une importance particulière. Nous vous en ferons la démonstration plus loin.

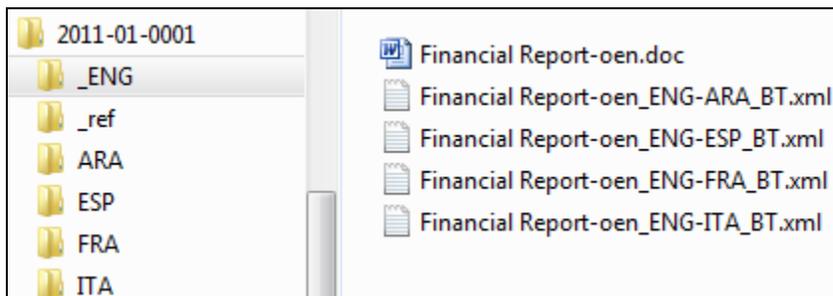
Au moment de créer les bitextes, AlignFactory donne le choix de créer le bitexte dans le répertoire du document source ou cible. Si vous choisissez de créer le bitexte dans le répertoire du document source, le bitexte portera le nom du document source. Sinon, le bitexte portera le nom du document cible.

Répertoire de création des bitextes :	
<input checked="" type="radio"/>	Répertoire des documents de la première langue
<input type="radio"/>	Répertoire des documents de la deuxième langue
<input type="radio"/>	Répertoire séparé :
<input type="text"/>	<input type="button" value="Parcourir..."/>

Voici deux exemples où le bitexte a été créé dans le répertoire du document cible.



Par défaut, la paire de langues du bitexte est identifiée dans le nom du bitexte, p. ex. ENG-FRA, ENG-ITA. La paire de langues dans le nom du bitexte n'est pas aussi fiable que le nom du fichier lui-même. C'est pour cette raison que nous vous recommandons d'utiliser des marqueurs précis dans le nom des documents. Dans ce cas-ci, le marqueur « o » n'apparaît pas dans le nom des bitextes, donc l'anglais est nécessairement le document original. Voici le cas inverse où les bitextes ont été créés dans le répertoire du document original.

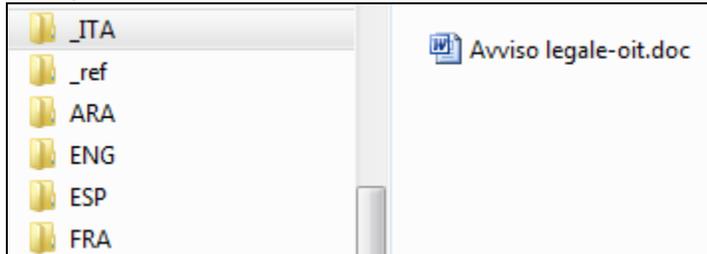


Le marqueur « oen » prouve que le document anglais est bien le document d'origine.

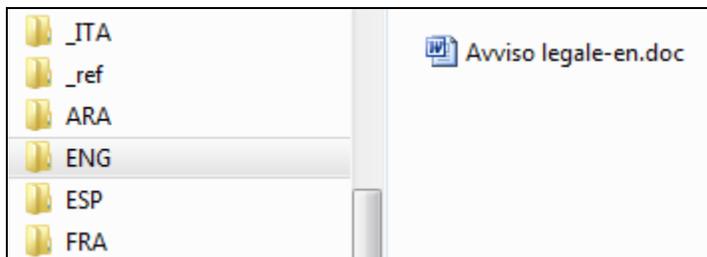
L'anglais comme langue cible

Pour l'exemple suivant, nous conserverons les mêmes paires de langue qu'aux exemples précédents. Admettons que le document source soit l'italien. Avant de traduire le document, nous ajoutons le suffixe « oit » au nom du document, ce qui signifie « original italien ».

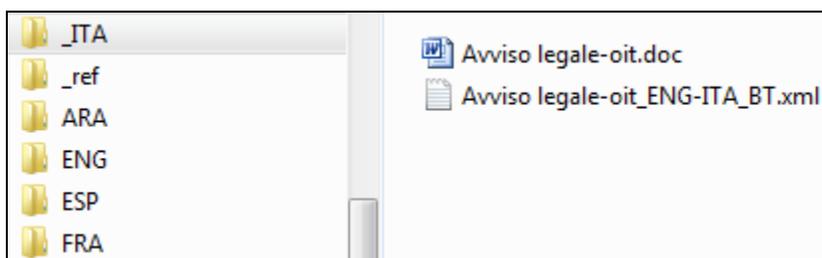
Exemple :



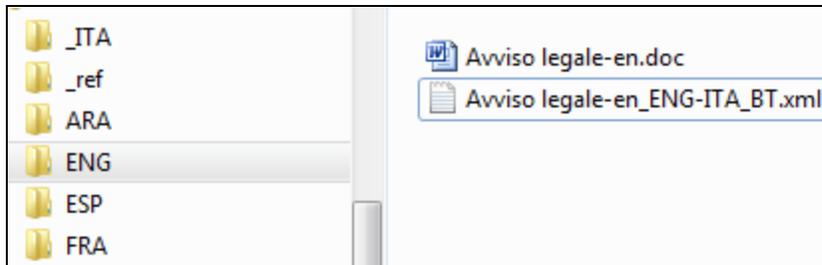
Si nous traduisons vers l'anglais, nous allons faire une copie de ce document dans le répertoire anglais et le nommer comme suit :



Une fois la traduction terminée, nous créons un bitexte. Si la paire de langue paramétrée dans AlignFactory est ENG-ITA, et que le bitexte est créé dans le répertoire cible, donc le répertoire ITA car l'italien est la deuxième langue dans la paire, voici comment le bitexte sera nommé :

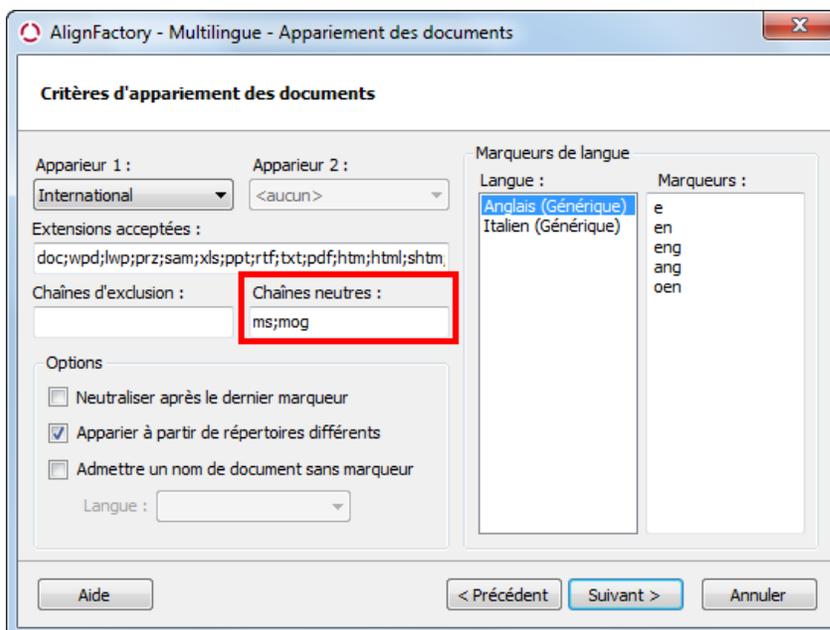


Si le bitexte est créé dans le répertoire de la langue source (la première langue de la paire paramétrée dans AlignFactory, l'anglais, et non pas celle du document source), voici comment le bitexte sera nommé :



Initiales dans le nom des documents

Si vous ajoutez des initiales dans le nom des documents, vous devez paramétrer AlignFactory afin qu'il puisse faire fi des initiales pour appairer les documents automatiquement. Le champ Chaînes neutres permet de neutraliser ce genre de marqueur.



Conclusion

Les exemples exposés dans le présent guide ne sont que des suggestions. Nous savons pertinemment qu'il existe d'autres méthodes efficaces. L'objectif principal est de vous fournir des lignes directrices sur la classification et le nommage de fichiers à partir desquels vous pouvez innover. Que vous possédiez un outil d'aide à la traduction ou non, ces méthodes demeurent néanmoins très pertinentes, car une partie de l'efficacité des employés en dépend directement. Cela dit, si vous avez des suggestions ou voulez nous faire part de votre méthode de classification ou de nommage de fichiers, nous serons ravis d'en prendre connaissance. Veuillez écrire à ventes@terminotix.com.

L'équipe Terminotix