



# Formation à l'utilisation professionnelle de Calames Partie Correspondant

## TP

*L'accès à la base de formation est garanti  
durant les 2 mois après la formation*

**Version du 25 novembre 2022**

## Sommaire des TP

TP 101 : Construire son arborescence .....	3
TP1 ÉLÉMENT DE CORRIGÉ : Construire son arborescence .....	5
TP 102 : Lire les traces pour diagnostiquer les problèmes de publication .....	8
TP102 CORRIGÉ : Lire les traces pour diagnostiquer les problèmes de publication .....	10
TP 103 : Export d'une branche de fichier EAD convertie en Dublin Core et au format tabulé.....	13
TP103 CORRIGÉ : Export d'une partie de fichier EAD convertie en Dublin Core au format tabulé.....	14

Les exercices de la formation partie Correspondant se font dans la continuité des exercices de la partie catalogueur

- Lancer Calames depuis le raccourci sur son bureau
- Choisir la base de Formation (à la place de « Calames »)
- S'identifier avec les mêmes identifiants communiqués par le formateur, construit sur le modèle RCRStagiairesXX (où XX est un nombre) et déjà utilisés pour la formation catalogueur.
- Les fichiers précédemment utilisés vont continuer de l'être

# TP 101 : Construire son arborescence

## Exercice 1 : comprendre une organisation

**Choisir un des deux exercices selon son niveau de maîtrise de la problématique : Ex 1.1 est le plus simple, l'Ex1.2 est le plus complexe des deux**

Dans l'interface publique de la base de production <http://www.calames.abes.fr/pub/>, en naviguant dans l'arborescence (« Accès direct » depuis le choix dans la liste « Bibliothèque(s) ») expliciter la logique d'organisation...

**TP101.Ex.1.1** ...des collections de l'EFA : ATHENES. Ecole française (jusqu'au niveau 3 sous le fichier maître pour certaines branches) :

**TP101.Ex.1.2** ...des collections du MNHN : PARIS. Muséum d'histoire naturelle (jusqu'au niveau 5 sous le fichier maître pour certaines branches) : exercice le plus complexe

NB : la fonction « Sommaire » accessible dans la colonne de droite, sous la boîte à outils peut faciliter le dépliement dans un seul affichage des différents niveaux

## Exercice 2 : définir liaison et inclusion

### TP101.Ex.2.1

**Importer le fichier** « Correspondance2.xml » (dossier TP101) en écrasant le fichier « RCRstagiairexx – ex nihilo ». Renommer ce fichier en « **RCRstagiaire xx –Correspondance2** »

### TP101.Ex.2.2

**Lier** les fichiers « Correspondance2 » et « Correspondance1 ».

### TP101.Ex.2.3

Reprendre le fichier « modèle » pour en faire le **fichier « maître »** : compléter les données EAD en vous limitant, pour l'exercice, à la reprise de l'élément indispensable au bon fonctionnement de Calames et au <unittitle> et renommer le fichier.

### TP101.Ex.2.4

**Inclure** le fichier « Correspondance1 », « ArchivesY » dans le fichier maître.

### TP101.Ex.2.5

Quelle correction/complément indispensable à faire (hors visio\_controle) avant de publier pour éviter de bloquer la publication ? Dans quel ordre doivent-ils être republiés ?



**Ne republier AUCUN des fichiers pour ne pas perturber le bon déroulement des TP suivants**

## Exercice 3 : définir une arborescence

Dans le dossier "TP101" [*chemin du répertoire fourni par les formateurs, clé usb, ou autre*], ouvrir le tableau « arborescence » : vous y trouvez un tableau, tel que l'Abes peut le fournir, d'une arborescence de fichiers EAD tel que présents dans la base de données Calames.

*L'exercice est fait collectivement au tableau*

### **TP101.Ex.3.1**

Y a-t-il des inclusions qui vous semble devoir être corrigées/modifiées ?

- quelles inclusions entre fichiers modifiez-vous ?
- pour quels fichiers modifiez-vous les ordres d'inclusions?

### **TP101.Ex.3.2**

Y a-t-il des fichiers qui mériteraient d'être scindés en 2 fichiers liés ? Lesquels ?

### **TP101.Ex.3.3**

Que pensez-vous de la cohérence de l'arborescence ?

Y a-t-il des intitulés à modifier pour une plus grande lisibilité ?

### **TP101.Ex.3.4**

Voyez-vous comment transmettre à l'Abes les demandes de modification des intitulés et inclusions ?

# TP1 ÉLÉMENT DE CORRIGÉ : Construire son arborescence

## Corrigé TP101-Exercice 1 : comprendre une organisation

### Ex1.1 : ATHENES. Ecole française

Sous fichier Maître (niv 0) : 3 séries relevant chacune d'un processus spécifique de production (niv 1)

- Fonds personnel de personnes
  - o en niveau 2 NB : pas de classement dans cette liste de fonds d'archéologues
- Production institutionnelle type des membres de l'Ecole : mémoires
  - o En niveau 2 classement chronologique
    - Niveau 3 : classement alphabétique
- Sites visités = donc classement thématique
  - o En niveau 2 classement alphabétique de noms de lieux mélangeant différents niveaux (pays, sites, régions, ...) ; ce n'est donc pas un classement géographique qui aurait aussi été possible

### Ex1.2 : PARIS. Muséum d'histoire naturelle

Sous le fichier maître (niv 0) : 5 sous maîtres selon une logique typologique de l'objet décrit

- Manuscrits et archives personnelles (niv. 1)
  - o En niveau 2 : organisé par lieu de conservation
- Archives institutionnelles (niv. 1)
  - o En niveau 2, voire 3 organisées en service producteur
- Oeuvres d'art et objets patrimoniaux (niv. 1)
  - o En niveau 2 poursuit la logique typologique par type d'objets d'art ; qui se poursuit au niveau 3, voir 4 et 5 ([Objets d'art déco.../Objets et dispositifs de médiation scientifique/ Modèles / Modèles en bois/ Modèles réduits d'apiculture](#) aux niveaux suivants on trouve
    - soit un classement topographique (par cote) : exemple [numismatique/moulages](#),
    - soit alphabétique : exemple [sculptures/statues](#) ; dans ce cas, on peut avoir un doute sur la nature des noms : sculpteur ou personne représentée => justifierait un <arrangement>
    - soit thématique ([Objets d'art déco.../Objet d'histoire.../Vues et images.../](#)
- Collections photographiques (niv. 1)
  - o En niveau 2 par photographes => par ordre alphabétique regroupés dans des sous-dossiers pour faciliter l'accès (éviter « effet râteau »)  
La sous-entrée « Par auteur » de niveau 3 ne se comprend que parce qu'il est prévu un frère « Photothèque du MNHN » non encore publié
- Instruments scientifiques : entrée non développée

## Corrigé TP101-Exercice 2 : Liaisons et inclusions

### **Liens vers le manuel et ses copies d'écran**

- <https://documentation.abes.fr/aidecalames/manueloutildecatalogage/index.html#OrganiserFichier>

### Ex2.1

Dans un cas réel dans la base de production, il faudrait reprendre en <archdesc> de « Correspondance2 » le contenu du <archdesc> de « Correspondance1 » pour que les héritages fonctionnent.

### Ex2.2

C'est bien dans les propriétés du fichier « Correspondance2 » que se paramètre la liaison au fichier « Correspondance1 »

### Ex2.5

Vérification à faire pour empêcher que la publication ne se bloque

- La présence du RCR Calames dans « Correspondance2 »
- L'absence de doublon dans ces fichiers

Ordre : attention à le faire dans l'ordre d'affichage public, à savoir :

- fichier maître, c'est-à-dire celui dans lequel un ou d'autres fichiers sont inclus,
- puis fichiers directement inclus dans le fichier maître
- puis fichier lié au précédent

Dans le cas présent il est possible de publier les différents fichiers

- soit dans cet ordre : fichier maître, « **correspondance1** », « **correspondance2** », « **archivesY** »
- soit dans celui-ci : fichier maître, « **correspondance1** », « **archivesY** », « **correspondance2** »,

## Corrigé TP101-Exercice 3 : définir une arborescence

### Ex3.1

Il y a un fichier (2458) hors arborescence dont il faut définir le parent (1370) et le n° d'ordre

Il y a une erreur dans les n° d'ordre des fichiers de niv 1, deux fois l'ordre « 4 » pour 1883 et 1611

Si cela ne pose actuellement pas de souci à l'affichage public c'est que l'un des 2 fichiers n'est pas publié

Il y a de multiples doublons d'ordre dans la fratrie enfant de 1370

### Ex3.2

Les fichiers 2471 (6,2 Mo) et 1376 (6 Mo) dépassent très largement la taille limite de 3 Mo, même avec la marge de tolérance (3,5 à 4 Mo)

### Ex3.3

Parmi la fratrie enfant de 1370, l'ordre alphabétique de nom n'est pas respecté

Les intitulés des fichiers 1376, 1662, 2146 et 2202 ne sont pas dans l'ordre nom\_prénom empêchant de reconstituer un ordre alphabétique avec l'intitulé

### Ex3.4

Pour transmettre les corrections à l'Abes il faut compléter les colonnes

Nouvelle clé du Parent = Insertion dans le document	Nouvelle pos = n° d'ordre	Nouvel intitulé (si besoin) :
---	---------------------------	-------------------------------

:

Pour régénérer un ordre alphabétique

- Copier la fratrie dans un nouvel onglet Excel
- La numéroter dans une colonne « ordre dans tableau »
- Ajouter 2 colonnes (issues de la matrice) « Nouvelle pos = n° d'ordre » et « Nouvel intitulé (si besoin) »
- Copier/coller les intitulés dans la colonne « Nouvel intitulé » et corriger ceux qui doivent l'être
- Trier le tableau sur ce Nouvel intitulé
- Générer un nouveau n° d'ordre de 10 en 10 dans la colonne « Nouvelle pos = n° d'ordre »
- Trier le tableau par « ordre dans tableau »
- Copier/coller le résultat des 2 colonnes « Nouvelle pos = n° d'ordre » et « Nouvel intitulé (si besoin) » dans le tableau source

# TP 102 : Lire les traces pour diagnostiquer les problèmes de publication

## Exercice 1 : lecture et compréhension des Traces de (dé)publication de fichiers

Pour chacun des fichiers traces au format txt qui se trouvent dans le dossier "TP102" [*chemin du répertoire fourni par les formateurs, clé usb, ou autre*], vous devez établir

- à quelle date et heure a été lancé le processus ? Sait-on à quelle heure il s'est achevé ?
- si la (dé)publication a aboutie ou pas
- si la (dé)publication n'a pas abouti, pour quelle raison ?
- Définir dans ce cas les mesures à prendre pour corriger ce qui a bloqué la (dé)publication

<b>TP102.Ex.1.1</b>	Cas du fichier « 50-Traces »
<b>TP102.Ex.1.2</b>	Cas du fichier « 51-Traces »
<b>TP102.Ex.1.3</b>	Cas du fichier « 309-Traces »
<b>TP102.Ex.1.4</b>	Cas du fichier « 1061-Traces »
<b>TP102.Ex.1.5</b>	Cas du fichier « 2185-Traces »
<b>TP102.Ex.1.6</b>	Cas du fichier « 3872-Traces »
<b>TP102.Ex.1.7</b>	Cas du fichier « 4609-Traces »



# TP102-Exercice 2 : Analyse pour publication du fichier « AbonnementZ »

## Situation

Vous aviez un fichier EAD importé dans Calames il y a longtemps que vous aviez tenté de publier.

Les traces d'alors signalaient

```
(...) ErrorMessage Analyse XML : ligne 783, caractère 74, valeur égale  
attendue [SQLSTATE 01000]  
ERROR GRAVE: la publication est abandonnée, le serveur de  
consultation reste en l'état antérieur [SQLSTATE 01000]  
Message 0, gravité 0, état 1 : Protocol error in TDS stream  
[SQLSTATE HY000]
```

## TP102.Ex.2.1

- Pour quelle raison le processus de publication a-t-il échoué ? quelle démarche devez-vous entreprendre ?

## TP102.Ex.2.2

Suite à une erreur de manipulation, vous avez supprimé le fichier de Calames. Heureusement vous aviez fait un export au format « Natif »

Dans le dossier "TP102" [*chemin du répertoire fourni par les formateurs, clé usb, ou autre*], vous trouvez cet export sous le nom « AbonnementZ »

- Modifier le fichier si besoin avant de le réimporter dans Calames sous une nouvelle clé et l'intitulé « RCRStagairexx – AbonnementZ »

NB : vous disposez

- de la Déclaration du format XML et de la DTD EAD au format txt dans le même dossier TP102
- de l'application Oxygen accessible sur le bureau de votre poste

# TP102 CORRIGÉ : Lire les traces pour diagnostiquer les problèmes de publication

## Corrigé TP102-Exercice 1 : lecture et compréhension des Traces de (dé)publication de fichiers EAD

### Ex1.1 : fichier 50

Les Traces indiquent qu'une opération a été lancée le 4 juillet 2022, vers 17h.  
Il s'agit d'une dépublication qui a abouti.

### Ex1.2 : fichier 51

Les Traces indiquent qu'une opération a été lancée le 26 juin 2009 à 22h10.  
Il s'agit d'une publication qui a abouti à 22h22.  
C'était une publication programmée le soir en-dehors des heures de grande activité et non immédiate.

### Ex1.3 : fichier 309

Les Traces indiquent qu'une opération a été lancée le 30 août 2010 vers 10h20.  
Il s'agit d'une dépublication qui a échoué.  
La raison de l'échec de cette dépublication est la trop grande sollicitation des ressources serveur hébergeant la base de données Calames et le processus automatique de priorisation de certaines opérations qui a désigné cette opération de dépublication comme « victime ».  
Dans ce cas, le fichier lui-même n'est pas en cause. Lorsqu'il se présente, il suffit de relancer la publication de préférence à un moment moins chargé (par la publication programmée).

### Ex1.4 : fichier 1061

Les Traces indiquent qu'une opération a été lancée le 22 novembre 2022 vers 16h.  
Il s'agit d'une publication qui a échoué.  
La raison de l'échec de cette publication est un non-respect du formalisme dans les données XML-EAD du fichier.  
Dans ce cas, il faut corriger l'erreur XML.

### Ex1.5 : fichier 2185

Les Traces indiquent qu'une opération a déjà été lancée le 26 juin 2021 vers 4h du matin.  
Il s'agit d'une publication qui a échoué.  
C'était une publication programmée et non immédiate.  
La raison de l'échec de cette publication est la trop grande sollicitation des ressources serveur. Ici plus spécifiquement, le système n'a pas réussi à copier les données depuis le serveur de production (celui sur lequel on travaille depuis l'interface professionnelle) vers le serveur (READ) qui sert à l'interface publique.  
Dans ce cas, le fichier lui-même n'est pas en cause. Lorsqu'il se présente, il suffit de relancer la publication de préférence à un moment moins chargé (par la publication programmée).

### Ex1.6 : fichier 3872

Les Traces indiquent qu'une opération a été lancée le 24 novembre 2022 à 16h.  
Il s'agit d'une publication partielle : en effet on voit outre la clé du fichier (3872), l'ID du composant (Calames-201583114830601) qui a échoué.  
La raison de l'échec de cette publication est un non-respect du formalisme dans les données XML-EAD du fichier.

*Dans ce cas, il faut corriger l'erreur XML*

**Ex1.7 : fichier 4609**

*Les Traces indiquent qu'une opération a été lancée le 24 novembre 2022 à 10h30.*

*Il s'agit d'une publication qui a échoué.*

*La raison de l'échec de cette publication est l'absence d'un RCR reconnu par Calames dans l'attribut AUTHFILENUMBER du <corpname> du <repository>.*

*Dans ce cas, il faut corriger les données dans Calames avant de relancer la publication.*

## Corrigé TP102-Exercice 2 : Cas du fichier « RCRstagiairexx – AbonnementZ »

### Ex2.1

Dans le cas de ce fichier, la publication a échoué en raison d'au moins une erreur XML dans le fichier. Cette erreur ne peut être corrigée dans Calames, il faut donc utiliser un éditeur XML extérieur à Calames. Habituellement, il faut alors exporter le fichier EAD au format « Natif-Déclaration ».

### Ex2.2

- Ouvrir « AbonnementZ » dans un éditeur XML (ici Oxygen)
- Ajouter avant la balise <ead> la déclaration disponible dans le dossier Tp102
- Lancer une vérification de conformité [à l'XML] et de validité [au regard de la DTD]. L'éditeur XML signale une première erreur. En cliquant sur le texte de l'erreur, vous accédez directement au bon endroit dans les données EAD.
- Faire la correction dans l'éditeur puis relancer la vérification.
- L'éditeur XML signale une deuxième erreur. La corriger. Enregistrer le fichier corrigé.
- Réimporter le fichier corrigé.  
Habituellement, cet import se fait en écrasant la version de départ.

#### Erreurs EAD à corriger dans Oxygen

Un nom de balise n'est pas conforme à la DTD EAD : balise < sujet > ligne 109 au lieu de < subject >  
< controlaccess >

```
< sujet subject source="Sudoc" authfilenumber="027805905" normal="Relations
internationales">Relations internationales</-sujet subject >
< geogname role="sujet" source="Sudoc" authfilenumber="027228231" normal="Grande-
Bretagne">Grande-Bretagne</geogname>
< geogname role="sujet" source="Sudoc" authfilenumber="02726369X" normal="Irlande"
>Irlande</geogname>
</controlaccess>
```

#### Erreur XML

Il y a du texte superflu dans la balise ouvrante qui est interprétée comme un attribut ligne 783 :

Le PCData a été par erreur saisi dans la balise ouvrante et non entre la balise ouvrante et la balise fermante.

```
< controlaccess >
< subject source="Sudoc" authfilenumber="027805905" normal="Relations internationales"
Relations internationales >-Relations internationales </subject >
< geogname role="sujet" source="Sudoc" authfilenumber="027228231" normal="Grande-
Bretagne">Grande-Bretagne</geogname >
< geogname role="sujet" source="Sudoc" authfilenumber="02726369X" normal="Irlande">
Irlande</geogname >
</controlaccess >
```

#### Erreur XML

Il y a une balise </c> fermante en trop avant </dsc>, à supprimer, ligne 16466

NB : selon le processus, la date d'import et la nature des erreurs XML, il est possible que vous trouviez dans Calames des fichiers non conformes à XML ou que le système refuse d'importer ces fichiers non conformes à l'XML.

# TP 103 : Export d'une branche de fichier EAD convertie en Dublin Core et au format tabulé

## TP103-Exercice 1 : Export DublinCore et son résultat en csv

Si cela n'a pas été déjà fait, importer le fichier « AbonnementZ », qui vient d'être corrigé lors du TP précédent. Une fois importé, l'exporter au format DublinCore

Ouvrir le résultat dans un tableau Excel

Si besoin la procédure est détaillée dans le manuel Correspondant de la documentation Calames.

### **Pour trouver la procédure :**

- Se rendre sur la [page d'accueil de ce manuel](#), puis en utilisant la recherche présente dans la colonne de gauche du manuel.

Si le fichier s'ouvre correctement dans Excel, on doit obtenir un tableur dans lequel chaque composant représente une ligne.

Répondre aux deux questions posées par les formateurs sur les données contenues dans le tableau une fois ouvert dans Excel.

## TP103-Exercice 2 : Export en Dublin Core avec XPath

Dans le fichier « AbonnementZ », on souhaite exporter toujours en Dublin Core au format csv seulement les 10 premiers <c> enfants du <c> « Début de la session 1880 », sans le <c> « Début de la session 1880 » lui-même.

Quel est le modèle d'export à utiliser ?

### **Pour construire l'XPath adapté :**

- Se rendre sur la [page d'accueil de ce manuel](#), puis en utilisant la recherche présente dans la colonne de gauche du manuel.

Pour vérifier que l'export a bien fonctionné, on peut ouvrir le fichier obtenu dans un bloc-notes pour regarder la fin du fichier (les derniers éléments EAD) et vérifier si tous les résultats voulus semblent y être. Le résultat est-il conforme aux composants attendus ?

## TP103 CORRIGÉ : Export d'une partie de fichier EAD convertie en Dublin Core au format tabulé

### Corrigé TP3-Exercice 1 : Export Dublin Core et son résultat en csv

Lien vers le manuel Correspondant pour l'ouverture d'un fichier csv dans Excel :

- <https://documentation.abes.fr/aidecalames/manuelcorrespondant/index.html#Tableur>

Si les données s'affichent correctement dans Excel, on doit obtenir les éléments Dublin Core en intitulés de colonnes, puis les 10 composants voulus avec 1 composant par ligne du tableur.

### Corrigé TP3-Exercice 2 : Export en Dublin Core avec XPath

Lien vers le manuel Correspondant pour trois exemples de construction de XPath :

[https://documentation.abes.fr/aidecalames/manuelcorrespondant/index.html#ExportModele\\_2](https://documentation.abes.fr/aidecalames/manuelcorrespondant/index.html#ExportModele_2)

L'axe d'XPath à utiliser pour sélectionner une partie des <c> d'une fratrie sans leur parent est l'axe **following-sibling::** ou **preceding-sibling::**

*Dans le cas de cet exercice, sélectionner les 10 premiers <c> de la fratrie implique d'utiliser l'axe preceding-sibling en se positionnant sur le 11<sup>e</sup> composant de la fratrie.*

L'XPath à construire prend donc la forme :

`//c[@id="ID du 11e <c> de la fratrie"]/preceding-sibling::c`

Attention les quotes encadrant l'ID du <c> sont des simples quotes et doivent être deux simples quotes, le filtre d'export Calames ne prenant pas en compte les doubles quotes.