



Mode opératoire d'utilisation de « BiblioMeter »

Contributeurs :

Amal Chabli (amal.chabli@orange.fr)

Ludovic Desmeuzes (ludovic.desmeuzes@yahoo.com)



Version du 2 juillet 2024

Table des matières

I.	Présentation de l'application	2
II.	Accès à l'application et à son exécutable	3
III.	Actions préalables incontournables	4
IV.	Lancement de l'application via l'exécutable	5
V.	Onglet « Analyse élémentaire des corpus »	9
V.1.	Section « Statut des fichiers de parsing »	9
V.2.	Section « Construction des fichiers de parsing par BDD »	10
V.3.	Section « Synthèse des fichiers parsing de toutes les BDD »	11
VI.	Onglet « Consolidation annuelle des corpus »	13
VI.1.	Étape « Croisement avec le fichier des effectifs LETI »	13
VI.2.	Étape « Résolution des homonymies »	15
VI.3.	Étape « Attribution des OTP »	15
VI.4.	Étape « Création de la liste consolidée des publications »	16
VII.	Onglet « Mise à jour des facteurs d'impact »	18
VII.1.	Mise à jour de la base de données des IFs	19
VII.2.	Mise à jour des IFs dans les listes consolidées	19
VIII.	Onglet « KPIs et graphes »	20
VIII.1.	Analyse des IFs et calcul des KPIs	20
VIII.2.	Analyse des collaborations	22
VIII.3.	Analyse des mots clefs	26
IX.	Attribution des OTPs	28
X.	Extraction d'un corpus de l'année	30
X.1.	Extraction pour l'Institut LITEN	30
X.2.	Extraction pour l'Institut LETI	32
XI.	Mise à jour du corpus de l'année	34
XII.	Complément HAL du corpus de l'année	34
XIII.	Mise à jour du fichier des effectifs de l'Institut	34
XIV.	Réduction du fichier « orphan.xlsx »	36
XIV.1.	La liste des auteurs et leurs affiliations sont ambiguës	36
XIV.2.	Les auteurs recherchés ont plusieurs affiliations dont celle de l'Institut	37
XIV.3.	Les auteurs recherchés n'ont qu'une seule affiliation, celle de l'Institut	37
XV.	Renseignement des fichiers des IFs manquants et des ISSN manquants	40
XVI.	Préparation annuelle de la base de données des IFs	42
XVII.	Choix de la période d'analyse	43
XVIII.	Arborescence du dossier « BiblioMeter_Files »	44

I. Présentation de l'application

BiblioMeter est une application logicielle dont l'objectif est de **faciliter le suivi de la production scientifique des instituts LITEN et LETI**. Elle utilise les données, qu'on appelle corpus, pouvant être extraites de **WoS et Scopus**, des plateformes d'information scientifique et technique sur lesquelles nous avons accès à des bases de données transdisciplinaires de résumés et de citations de publications scientifiques.

Elle a été développée en langage « python » par Ludovic Desmeuzes (ludovic.desmeuzes@yahoo.com) dans le cadre d'une formation en alternance d'un mastère spécialisé BIG DATA de GEM et ENSIMAG de 2021 à 2022, encadré par Amal Chabli (amal.chabli@orange.fr) et François Bertin (francois.bertin7@wanadoo.fr) et selon le cahier des charges spécifié par Jean-Pierre Simonato (jean-pierre.simonato@cea.fr) puis complété par Jean-Paul Barnes (jean_paul_barnes@cea.fr) et Christel Faure (christel.faure@cea.fr) lors de l'extension de l'application à l'Institut Leti. Elle est basée sur l'utilisation des bibliothèques de fonctions de 2 packages « **bmfuncts** » et « **bmgui** » déposés sur GitHub et accessibles via le lien : <https://github.com/TickyWill/BiblioMeter.git>

Elle utilise les fonctions d'analyse élémentaire des corpus du package « BiblioParsing » développé également en langage python par François Bertin et Amal Chabli selon le cahier des charges spécifié par Iona Brenac (iona.brenac@gmail.com). Ce package est également disponible sur GitHub (<https://github.com/TickyWill/BiblioParsing>).

L'utilisation de l'application peut se faire via un exécutable et dans ce cas elle ne requière pas l'installation de « python » ni l'accès à la Toile. **Les informations sensibles utiles à l'application et les résultats des traitements sont présents uniquement sur le réseau interne du CEA** et requiert donc soit une connexion filaire soit une connexion VPN au réseau du CEA via « FortiClient ».

La version en service de l'application est la **version 4.0.0**.

II. Accès à l'application et à son exécutable

Pour utiliser l'application, copier sur votre bureau le **raccourci de l'exécutable** disponible dans le dossier « \BiblioMeter » qui est accessible depuis un **PC connecté au réseau du CEA** (filaire ou VPN).

En date du **02/07/2024**, l'**exécutable en service** est lancé via le raccourci nommé :

« 2024_07_02 BiblioMeter.exe – Raccourci »
--

Le **dossier de travail** de cet exécutable se trouve également dans le dossier « \BiblioMeter » et il est nommé :

« BiblioMeter_Files »

Pour le LITEN, le dossier « \BiblioMeter » est accessible via le chemin suivant :

[\\\(Katmandou\)](#) (S:)

→ :\130-LITEN

→ \130.1-Direction

→ \130.1.2-Direction Scientifique

→ \130.1.2.2-Infos communes

→ \BiblioMeter

Une **sauvegarde du dossier de travail** nommée « BiblioMeter_Files_sauvegarde » est gérée par l'équipe de Direction via le chemin suivant :

[\\\(Katmandou\)](#) (S:)

→ :\130-LITEN

→ \130.1-Direction

→ \130.1.2-Direction Scientifique

→ \130.1.2.1-Dossiers en cours

→ \03- Publications

→ \BiblioMeter

Le dossier « \BiblioMeter » contient également les raccourcis vers les **éventuelles anciennes versions** de l'exécutable dont le dossier de travail est renommé « BiblioMeter_Files_aaaa-mm-jj » avec « aaaa-mm-jj » la date caractérisant l'exécutable associé.

Pour le LETI, le dossier « \BiblioMeter » est accessible via le chemin suivant :

[\\\(Katmandou\)](#) (S:)

→ :\120-LETI

→ \120.38-BiblioMeter

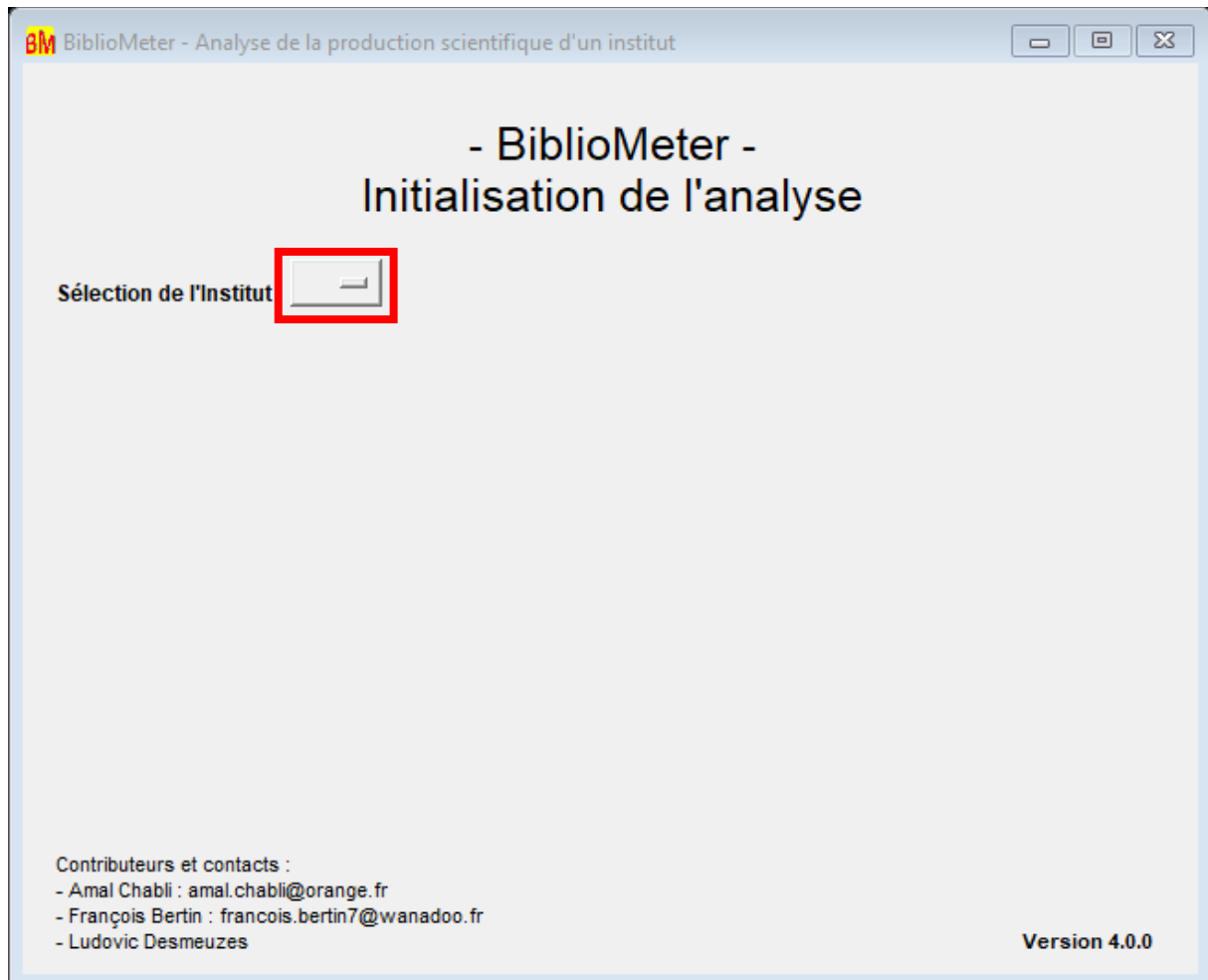
III. Actions préalables incontournables

Avant l'utilisation de l'application, il convient de vérifier le statut (effectuées ou non) des actions de la séquence générale décrite dans le tableau ci-dessous. Ce tableau donne la périodicité et le lien vers le mode opératoire de chacune des actions.

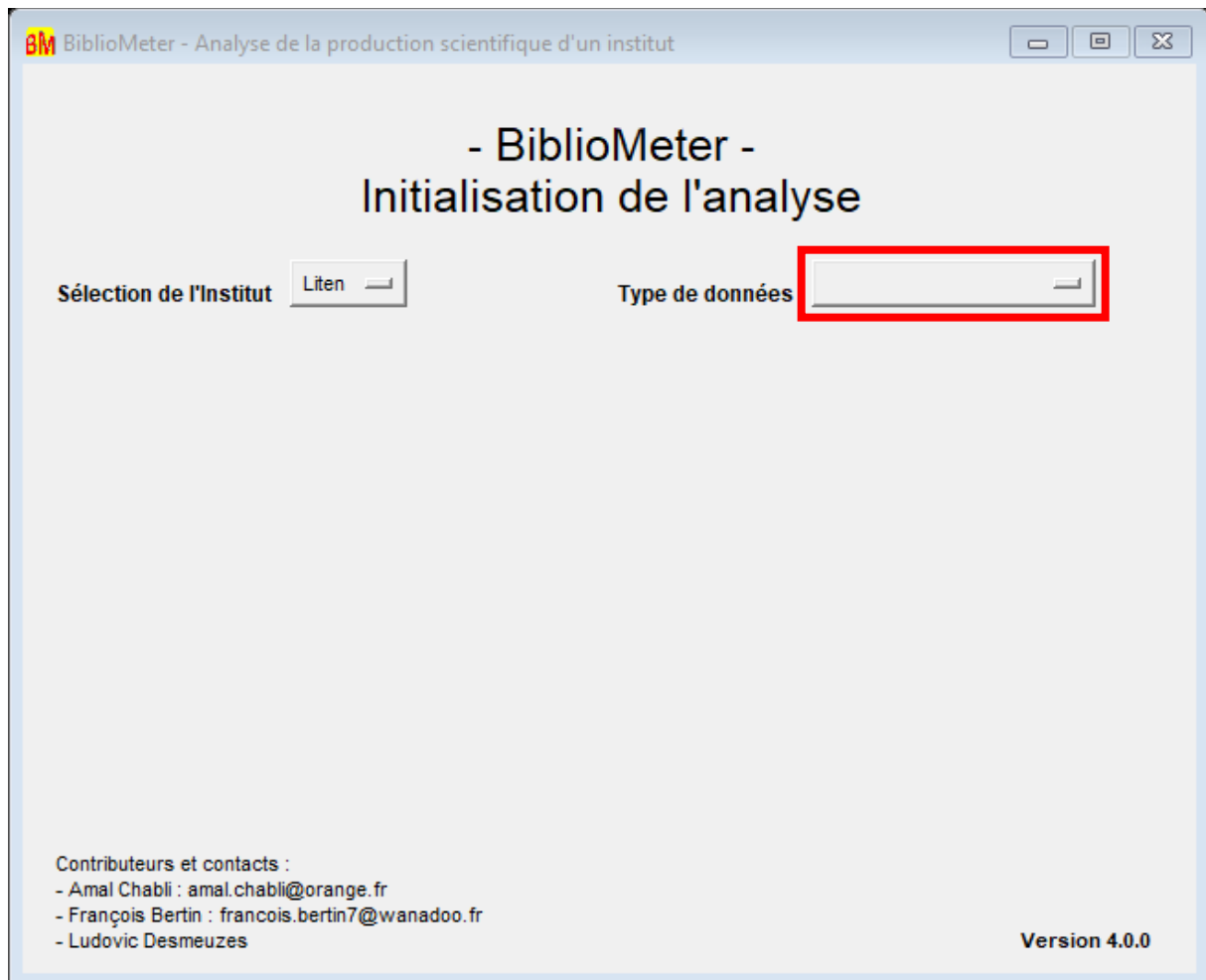
Action	Mode opératoire	Périodicité
1. Création d'un dossier pour l'année à analyser, année <n>	Arborescence du dossier « \BiblioMeter_Files » Partie XVIII	Annuelle Année <n>/Janvier
2. Extraction des corpus bruts de Scopus et WoS de l'année <n>	Extraction du corpus de l'année Partie X	Selon l'Institut
3. Mise à jour des corpus bruts de Scopus et WoS de l'année <n>	Mise à jour du corpus de l'année Partie XI	Annuelle Année <n+1>/Juillet
4. Complément du corpus brut de Scopus par HAL de l'année <n>	Complément HAL du corpus de l'année Partie XII	Annuelle Année <n+1>/Juillet
5. Mise à jour du fichier des effectifs de l'Institut	Mise à jour du fichier des effectifs de l'Institut Partie XIII	Trimestrielle
6. Préparation du fichier des facteurs d'impact de l'année <n-1>	Préparation de la base de données des IFs Partie XVI	Annuelle Année <n>/Juillet
7. Lancement de l'application	Lancement de l'application via un exécutable Partie IV	À volonté , si 1 à 6 vérifiées

IV. Lancement de l'application via l'exécutable

1. **Double cliquer sur l'exécutable** et attendre que la fenêtre de lancement s'ouvre en proposant la sélection de l'Institut concerné par l'analyse comme ci-dessous.



2. **Sélectionner l'Institut dans la liste déroulante** qui s'affiche en cliquant sur le bouton encadré dans la fenêtre ci-dessus et attendre que la fenêtre se mette à jour comme ci-dessous.



3. **Sélectionner le type de données à traiter dans la liste déroulante** qui s'affiche en cliquant sur le bouton encadré dans la fenêtre ci-dessus et attendre que la fenêtre se mette à jour comme ci-dessous. Le choix se fait entre les 3 types de données suivantes :
- « **Scopus & Wos** » : la combinaison de WoS avec Scopus ;
 - « **Scopus-Hal & Wos** » : la combinaison de WoS avec Scopus complété de HAL ;
 - « **Wos** » : l'utilisation de WoS seul.

Ce choix va se traduire par la mise en place des fichiers ad-hoc des données brutes dans le dossier « **.../Corpus/<database>/rawdata** » ou <database> est soit « scopus » soit « wos ».

BiblioMeter - Analyse de la production scientifique d'un institut

- BiblioMeter - Initialisation de l'analyse

Sélection de l'Institut Liten Type de données Scopus & WoS

Dossier de travail S:\130-LITEN\...\130.1.2.2-Infos communes\BiblioMeter\BiblioMeter_Files
Changer de dossier de travail

Liste des corpus ['2019', '2020', '2021', '2022', '2023', '2024']
Créer un nouveau dossier de corpus annuel

Lancer l'analyse

Contributeurs et contacts :
- Amal Chabli : amal.chabli@orange.fr
- François Bertin : francois.bertin7@wanadoo.fr
- Ludovic Desmeuzes

Version 4.0.0

4. **Vérifier le chemin du « Dossier de travail ».** Il doit pointer bien vers le dossier « \BiblioMeter_Files » qui est disponible dans le dossier « \BiblioMeter » accessible par le chemin précisé au paragraphe II à partir duquel l'application va travailler.
5. Si ce n'est pas le cas, **cliquer sur le bouton « Changer de dossier » (cadre rouge)** et **changer le chemin d'accès à « \BiblioMeter_Files »** avec l'aide de l'explorateur de fichiers qui s'ouvre.
6. Vérifier que le chemin pointe effectivement vers un dossier « \BiblioMeter_Files » comme dans l'exemple de la fenêtre ci-dessus.
7. Au premier traitement d'une nouvelle année ##### (ex : 2025), créer un nouveau dossier de corpus pour cette année en cliquant sur le bouton « **Créer un nouveau dossier de corpus annuel** » (cadre vert). Cette opération mettra en place, dans le dossier de travail, un dossier nommé « ##### » avec l'arborescence requise. Avant de poursuivre l'analyse, il convient de disposer des extractions Scopus et WoS correspondantes en suivant la procédure précisée au paragraphe X.
8. **Cliquez sur le bouton « Lancer l'analyse ».** Une fenêtre s'ouvre et propose 4 onglets encadrés dans l'exemple de la fenêtre ci-dessous et intitulés :
 - Analyse élémentaire des corpus
 - Consolidation annuelle des corpus
 - Mise à jour des facteurs d'impact
 - KPIs et graphes

Les opérations liées à chacun de ces 3 onglets sont décrites dans les sections suivantes de ce document. Notons que **l'exécution de l'application peut être interrompue à tout moment** pour reprendre les traitements ultérieurement sans perdre les résultats des traitements effectués si on ne change pas de type de données. Ainsi, chacun des 4 onglets propose un bouton « **Quitter** » à cet effet.

BiblioMeter - Analyse de la production scientifique d'un institut

Analyse élémentaire des corpus | Consolidation annuelle des corpus | Mise à jour des facteurs d'impact | KPIs et graphes

Analyse élémentaire des corpus du Liten

Données Scopus & WoS

Statut des fichiers de Parsing	Wos Données brutes	Wos Parsing	Scopus Données brutes	Scopus Parsing	Synthèse du parsing des BDD
Année 2019	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Année 2020	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Année 2021	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Année 2022	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Année 2023	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Année 2024	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

Mettre à jour le statut des fichiers

Construction des fichiers de Parsing par BDD

Sélection de l'année: 2024 | Sélection de la BDD: scopus | Lancer le Parsing

Synthèse des fichiers de Parsing de toutes les BDD

Sélection de l'année: 2024 | Lancer la synthèse | Quitter

V. Onglet « Analyse élémentaire des corpus »

Cette section de l'application sert à redistribuer par catégories (« parser ») les informations extraites des bases de données WoS et Scopus. À cette occasion, elles sont mises sous un format standardisé (indépendant de la base d'origine) et exploitable (gestion des caractères spéciaux), puis concaténées pour supprimer les doublons entre bases, pour qu'elles puissent être utilisées par les opérations des 3 autres onglets.

À la sélection de cet onglet le contenu de la fenêtre est équivalent à celui de l'exemple de fenêtre précédent. Les différents éléments qui la composent sont décrits ci-après.

V.1. Section « Statut des fichiers de parsing »

Le statut des fichiers dans le dossier « \BiblioMeter_Files » issus des différentes bases de données est affiché :

- Une **case cochée** indique que le fichier est présent
- Une **case non cochée** indique l'absence du fichier.

Le tableau se met automatiquement à jour après chaque action dans cet onglet, mais aussi en cliquant sur le bouton « **Mettre à jour le statut des fichiers** ».

Dans l'exemple de la fenêtre ci-dessous, toutes les opérations de « parsing » ont été effectuées pour les années 2019 à 2023, mais des cases sont non cochées pour l'année 2024.

BiblioMeter - Analyse de la production scientifique d'un institut

Analyse élémentaire des corpus | Consolidation annuelle des corpus | Mise à jour des facteurs d'impact | KPIs et graphes

Analyse élémentaire des corpus du Liten

Données Scopus & WoS

Statut des fichiers de Parsing	Wos Données brutes	Wos Parsing	Scopus Données brutes	Scopus Parsing	Synthèse du parsing des BDD
Année 2019	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Année 2020	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Année 2021	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Année 2022	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Année 2023	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Année 2024	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Mettre à jour le statut des fichiers

Construction des fichiers de Parsing par BDD

Sélection de l'année: 2024 | Sélection de la BDD: scopus | Lancer le Parsing

Synthèse des fichiers de Parsing de toutes les BDD

Sélection de l'année: 2024 | Lancer la synthèse | Quitter

Pour pouvoir passer aux onglets suivants, toutes les cases doivent être cochées lors de la mise à jour du statut des fichiers. Dans ce but et pour l'année à compléter :

- a. Les **données brutes des bases de données** doivent être rendues disponibles dans le dossier « **\BiblioMeter_Files** » (cf. § III action 2) ;
- b. Le « parsing » des données brutes doit être effectué en utilisant la section « **Construction des fichiers de parsing par BDD** » (cf § V.2 ;
- c. La synthèse du « parsing » des différentes bases de données doit être effectuée en utilisant la section « **Synthèse des fichiers parsing de toutes les BDD** » (cf § V.3).

V.2. Section « Construction des fichiers de parsing par BDD »

Cette section permet d'effectuer le « parsing » des données brutes pour chacune des bases de données de l'année à traiter par les opérations suivantes :

1. Sélectionner l'année à traiter par le menu déroulant en face de « **Sélection de l'année** » ;
2. Sélectionner le type de base de données (Scopus ou WoS) par le menu déroulant en face de « **Sélection de la BDD :** »
3. Puis cliquer sur le bouton « **Lancer le parsing** ».

L'application va effectuer le « parsing », afficher un message pour indiquer que le « parsing » a été effectué et mettre à jour automatiquement le tableau du statut des fichiers.

Cette action a pour effet de créer des fichiers avec l'extension « .dat » dans le dossier « **\BiblioMeter_Files\<année>\Corpus\<BDD>\parsing** » où <année> désigne l'année traitée au format « aaaa » et <BDD> désigne soit WoS soit Scopus.

L'exemple ci-dessous montre l'évolution de la fenêtre suite à la construction des fichiers de « parsing » pour l'année 2024 et la base de données Scopus après avoir rendu disponible les données brutes de cette base.

BiblioMeter - Analyse de la production scientifique d'un institut

Analyse élémentaire des corpus | Consolidation annuelle des corpus | Mise à jour des facteurs d'impact | KPIs et graphes

Analyse élémentaire des corpus du Liten

Données Scopus & WoS

Statut des fichiers de Parsing	Wos Données brutes	Wos Parsing	Scopus Données brutes	Scopus Parsing	Synthèse du parsing des BDD
Année 2019	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Année 2020	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Année 2021	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Année 2022	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Année 2023	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Année 2024	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Mettre à jour le statut des fichiers

Construction des fichiers de Parsing par BDD

Sélection de l'année: 2024 | Sélection de la BDD: scopus | Lancer le Parsing

Synthèse des fichiers de Parsing de toutes les BDD

Sélection de l'année: 2024 | Lancer la synthèse | Quitter

V.3. Section « Synthèse des fichiers *parsing* de toutes les BDD »


Cette section permet d'effectuer la synthèse des fichiers issus du « parsing » des bases de données WoS et Scopus par les opérations suivantes :

1. Sélectionner l'année à traiter par le menu déroulant en face de « **Sélection de l'année** » ;
2. Puis cliquer sur le bouton « **Lancer la synthèse** ».

L'application va effectuer la synthèse, afficher un message pour indiquer que la synthèse a été effectuée et mettre à jour automatiquement le tableau du statut des fichiers. Cette étape de synthèse ne peut être effectuée que si les fichiers de « parsing » des bases de données WoS et Scopus sont disponibles.

Cette action a pour effet de créer (en particulier) des fichiers avec l'extension « .dat » dans le dossier « \BiblioMeter_Files\<année>\Corpus\deduplication\parsing » où <année> représente l'année traitée au format « aaaa ».

L'exemple ci-dessous montre l'évolution de la fenêtre suite à la synthèse des fichiers « parsing » de toutes les BDD pour l'année 2024 une fois le parsing des bases de données WoS et Scopus disponibles.


Bibliometer - Analyse de la production scientifique d'un institut

Analyse élémentaire des corpus
Consolidation annuelle des corpus
Mise à jour des facteurs d'impact
KPIs et graphes

Analyse élémentaire des corpus du Liten

Données Scopus & WoS

Statut des fichiers de Parsing	Wos Données brutes	Wos Parsing	Scopus Données brutes	Scopus Parsing	Synthèse du parsing des BDD
Année 2019	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Année 2020	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Année 2021	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Année 2022	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Année 2023	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Année 2024	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

Mettre à jour le statut des fichiers

Construction des fichiers de Parsing par BDD

Sélection de l'année
2024
Sélection de la BDD
scopus
Lancer le Parsing

Synthèse des fichiers de Parsing de toutes les BDD

Sélection de l'année
2024
Lancer la synthèse

Quitter

Le traitement peut passer maintenant à la consolidation annuelle des corpus en cliquant sur l'onglet éponyme.

VI. Onglet « Consolidation annuelle des corpus »

Cette section de l'application utilise les fichiers créés dans l'onglet « **Analyse élémentaire des corpus** » pour consolider le corpus de chaque année en 4 étapes successives qui sont :

1. Croisement auteurs-effectifs de l'Institut
2. Résolution des homonymies
3. Attribution des OTP
4. Création de la liste consolidée des publications

À la sélection de cet onglet le contenu de la fenêtre est équivalent à celui de l'exemple de fenêtre ci-dessous. Les 4 étapes qui la composent sont décrites ci-après et sont à effectuer pour l'année sélectionnée par le menu déroulant en face de « **Sélection de l'année** ». L'année peut être modifiée après chacune des 4 étapes pourvu que pour chaque année **l'ordre d'exécution des étapes soit respecté**. L'exécution des étapes est obtenue par l'utilisation du bouton correspondant.

BiblioMeter - Analyse de la production scientifique d'un institut

Analyse élémentaire des corpus | Consolidation annuelle des corpus | Mise à jour des facteurs d'impact | KPIs et graphes

Consolidation annuelle des corpus du Liten

Données Scopus & WoS

Sélection de l'année: 2024

Etape 1 : Croisement auteurs-effectifs de l'institut

☐ Mettre à jour les effectifs de l'institut avant le croisement (coché = OUI) ?

Effectuer le croisement auteurs-effectifs

Etape 2 : Résolution des homonymies

Créer le fichier pour la résolution des homonymies

Etape 3 : Attribution des OTPs

Créer les fichiers pour l'attribution des OTPs

Etape 4 : Consolidation de la liste des publications

Créer la liste consolidée des publications

Quitter

VI.1. Étape « Croisement avec le fichier des effectifs LETI »

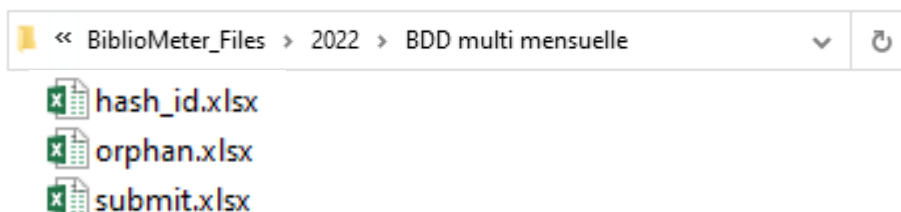
Cette première étape permet de croiser le corpus de l'année construit dans l'onglet « Analyse élémentaire des corpus » avec le fichier des effectifs de l'Institut. L'objectif de ce croisement est de recueillir les **informations utiles relatives aux auteurs** des publications **dont l'affiliation est l'Institut**. Il s'agit en particulier d'**identifier leur département** de rattachement.

Le fichier des effectifs de l'Institut peut être mis à jour lorsque des informations complémentaires sont fournies par le service RH de l'Institut en cochant la case en face de « **Mettre à jour les effectifs de l'Institut avant le croisement (coché = OUI) ?** ». Consultez l'opération « **Mise à jour du fichier des effectifs de l'Institut** » (cf § II Erreur ! Source du renvoi introuvable. action 3) pour préparer cette mise à jour avec le fichier d'informations complémentaires fournies par le service RH de l'Institut.

Le croisement s'effectue par les opérations suivantes :

1. Sélectionner l'année à traiter par le menu déroulant en face de « **Sélection de l'année** » ;
2. Cocher ou non la case pour mettre à jour ou non le fichier des effectifs de l'Institut ;
3. Puis cliquer sur le bouton « **Effectuer le croisement auteurs-effectifs** ». Cette opération prend entre 1 et 2 minutes et un message informe de sa complétion ou des erreurs rencontrées.

Cette opération a pour effet de créer plusieurs fichiers avec l'extension « .xlsx » dans le dossier « **\BiblioMeter_Files\<année>\0 - BDD multi mensuelle** » où <année> désigne l'année traitée au format « aaaa ». Le contenu de ce dossier pour l'année 2022 est donné en exemple ci-dessous.



Le fichier « **submit.xlsx** » est la base de données complète qui comporte **les publications avec les informations sur les auteurs affiliés à l'Institut**. Ce fichier de structure complexe est délicat à utiliser pour une analyse de la production scientifique. **La suite de l'application permet de rationaliser le contenu de ce fichier pour une analyse effective.**

Néanmoins, pour l'utiliser comme base d'analyse, il est recommandé de **travailler sur une copie** pour éviter d'avoir à refaire l'opération de croisement.

Le fichier « **hash_id** » contient **l'identifiant universel de chaque publication** construit à partir du titre, du premier auteur, du DOI, de l'ISSN et de l'année de première publication. Il est indépendant de l'extraction des bases de données WoS et Scopus.

Le fichier « **orphan.xlsx** » contient la liste des publications dont un auteur ayant l'affiliation de l'Institut n'a pas été trouvé dans le fichier des effectifs de l'Institut. L'utilisation de ce fichier est basée sur une recherche manuelle des raisons de cette absence. Cette recherche permet de compléter les fichiers présents dans le dossier « **\BiblioMeter_Files\4–Traitement Orphan** ». Ces fichiers sont utilisés pour corriger/compléter le croisement auteurs-effectifs :

- a. Suivez les instructions de l'opération « **Réduction du fichier « orphan.xlsx** » » (cf § XIV) pour utiliser le fichier « **orphan.xlsx** » ;
- b. Puis, lancer à nouveau le croisement en cliquant sur le bouton « **Effectuer le croisement auteurs-effectifs** ». Cette opération prend entre 1 et 2 minutes et un message informe de sa complétion ou des erreurs rencontrées.
- c. Répéter ces 2 opérations jusqu'à aboutir à un fichier « **orphan.xlsx** » vide. Cela indique que tous les auteurs vraiment affiliés à l'Institut ont été identifiés.

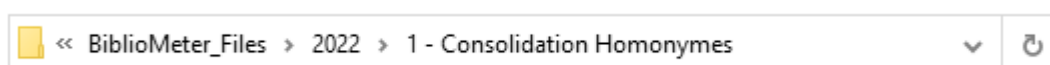
VI.2. Étape « Résolution des homonymies »

Cette deuxième étape permet de créer un fichier qui met en évidence les auteurs homonymes de l'Institut. Les homonymes sont définis ici comme les auteurs qui possèdent le même nom de famille et la même initiale (ou les mêmes initiales) de prénom.

La création de ce fichier s'effectue par les opérations suivantes :

1. Sélectionner l'année à traiter par le menu déroulant en face de « **Sélection de l'année** » ;
2. Puis cliquer sur le bouton « **Créer le fichier pour la résolution des homonymies** ». Cette opération est instantanée et un message informe de sa complétion ou des erreurs rencontrées.

Le fichier est créé dans le dossier « \BiblioMeter_Files\<année>\1 – Consolidation homonymes » où <année> indique l'année traitée. L'exemple ci-dessous est pour l'année 2022.



Il est nommé « **Fichier Consolidation <année>.xlsx** » où <année> indique l'année traitée. Dans ce fichier les homonymes sont indiqués par les cellules surlignées en jaune. Ces cellules contiennent les identifiants complets issus du fichier des effectifs de l'Institut pouvant correspondre au même nom et à la même initiale (ou aux mêmes initiales) de prénom.

Dans l'exemple ci-dessous, « Arnaud Jay » et « Amandine Jay » sont homonymes parce qu'identifiés comme auteurs par « A. Jay ».

2	62	2 Integratio energy an Merlet Y	10.1016/j.0378-7788 Article	224189 JAY	ARNAUD	DTS	SIRE	LELA	HOMONYMERLET, YANNIS; JAY, ARNAUD; JAY, AMANDINE
3	62	2 Integratio energy an Merlet Y	10.1016/j.0378-7788 Article	265848 JAY	AMANDINE	DTCH	STH2	['LES', 'LCS	HOMONYMERLET, YANNIS; JAY, ARNAUD; JAY, AMANDINE

La résolution de ces homonymies nécessite une intervention humaine via les opérations suivantes :

1. Ouvrir le fichier « **Fichier Consolidation <année>.xlsx** » sous EXCEL ;
2. Pour chacun des cas d'homonymie déterminer la personne de l'Institut qui est effectivement l'auteur de la publication ;
3. Supprimer la ligne correspondant à l'homonyme qui n'est pas auteur de la publication ;
4. Sauvegarder le fichier.

Dans l'exemple ci-dessus, c'est « Arnaud Jay » l'auteur de la publication, et non « Amandine Jay ». **La ligne entière d'« Amandine Jay » doit être supprimée** : sélectionnez la totalité de la ligne, cliquez droit puis « Supprimer ». Le résultat dans le fichier sauvegardé est alors le suivant :

2	62	2 Integratio energy an Merlet Y	10.1016/j.0378-7788 Article	224189 JAY	ARNAUD	DTS	SIRE	LELA	HOMONYMERLET, YANNIS; JAY, ARNAUD; JAY, AMANDINE
3	63	2 Integratio energy an Merlet Y	10.1016/j.0378-7788 Article	265848 JAY	AMANDINE	DTCH	STH2	['LES', 'LCS	HOMONYMERLET, YANNIS; JAY, ARNAUD; JAY, AMANDINE

La création du fichier pour la résolution des homonymies et **la résolution elle-même sont nécessaires à la poursuite du traitement.**

VI.3. Étape « Attribution des OTP »

Cette troisième étape permet de créer **un fichier EXCEL par département** contenant la liste des publications dont le premier auteur de l'Institut est affilié à ce département. Une colonne est disponible pour sélectionner manuellement l'OTP de chaque publication.

En même temps, le matricule des homonymes retenus comme auteurs à l'étape précédente sont conservés dans le fichier « Homonymes conservés » sauvegardé dans le dossier « \BiblioMeter_Files\<année>\4 – Informations », où <année> indique l'année traitée, en

correspondance avec l'identifiant universel des publications et serviront à résoudre automatiquement les homonymies lors d'un nouveau traitement du corpus de cette année.

La création des fichiers pour l'attribution des OTPs s'effectue par les opérations suivantes :

1. Sélectionner l'année à traiter par le menu déroulant en face de « **Sélection de l'année** » ;
2. Puis cliquer sur le bouton « **Créer les fichiers pour l'attribution des OTPs** ». Cette opération est instantanée et un message informe de sa complétion ou des erreurs rencontrées.

Les fichiers sont créés dans le dossier « **\BiblioMeter_Files\<année>\2 – OTP** » où <année> indique l'année traitée. Pour le **mode opératoire** d'attribution des OTPs suivre les instructions du paragraphe **IX**.

La création des fichiers pour l'attribution des OTPs est nécessaire à la poursuite du traitement mais **l'attribution des OTPs elle-même est facultative**.

VI.4. Étape « Création de la liste consolidée des publications »

Cette quatrième étape permet de créer un fichier qui prend en compte les informations renseignées dans les deux précédentes étapes. C'est-à-dire la résolution des homonymies et l'attribution des OTPs.

En même temps, les fichiers suivants sont créés ou mis à jour :

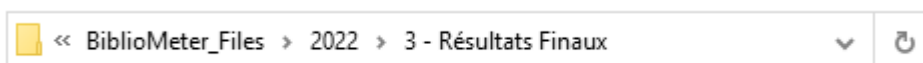
- Les OTPs attribués à l'étape précédente sont conservés dans le fichier « **OTPs conservés** » sauvegardé dans le dossier « **\BiblioMeter_Files\<année>\4 – Informations** », où <année> indique l'année traitée, en correspondance avec l'identifiant universel des publications et seront automatiquement attribués lors d'un nouveau traitement du corpus de l'année.
- Les facteurs d'impact IFs non connus dans la base de données des IFs sont disponibles dans le fichier « **<année>_IF manquants.xlsx** » sauvegardé dans le dossier « **\BiblioMeter_Files\<année>\3 – Résultats Finaux** » où <année> indique l'année traitée.
- Les journaux dont les ISSNs de la source (Scopus ou WoS par exemple) ne sont pas retrouvés dans la base de données des IFs sont disponibles dans le fichier « **<année>_ISSN manquants.xlsx** » sauvegardé dans le dossier « **\BiblioMeter_Files\<année>\3 – Résultats Finaux** » où <année> indique l'année traitée.

Ces 2 derniers fichiers sont utilisés par la mise à jour de la base de données des IFs dans l'onglet « **Mise à jour des IFs** ».

La création de la liste consolidée des publications s'effectue par les opérations suivantes :

1. Sélectionner l'année à traiter par le menu déroulant en face de « **Sélection de l'année** » ;
2. Cliquer sur le bouton « **Créer la liste consolidée des publications** ». Cette opération peut prendre 1 à 2 minutes et un message informe de sa complétion ou des erreurs rencontrées.

Le fichier est créé dans le dossier « **\BiblioMeter_Files\<année>\3 – Résultats Finaux** » où <année> indique l'année traitée. L'exemple ci-dessous est pour l'année 2022.

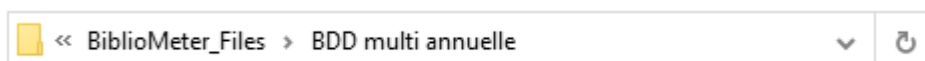


Il est nommé « **Liste consolidée aaaa.xlsx** ». C'est un document qui fait la synthèse des documents créés pour chaque département de l'Institut et éventuellement modifiés par l'attribution des OTPs à l'étape précédente (cf § **VI.3**).

De plus, cette liste consolidée est décomposée en plusieurs listes spécifiques des types de publications. Les types de publications définis actuellement sont les suivants :


- « Articles & Proceedings » : articles de journaux et communications d'actes de conférences indexés dans les bases de données,
- « Books & Editorials » : ouvrages, chapitres d'ouvrage et éditoriaux ;
- « Others » : tous ceux qui n'ont pas pu être classés dans les listes précédentes du fait que leur type de document est nouveau et doit être ajouté par l'administrateur de l'application dans la définition du paramètre global « DOC_TYPE_DICT » du module « BM_PubGlobal.py » de l'application.

Enfin, cette liste consolidée est concaténée avec les listes consolidées des années précédentes qui sont disponibles. Cette concaténation est sauvegardée dans le dossier suivant :

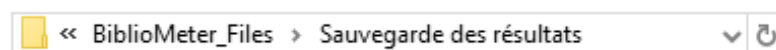


Le fichier correspondant à cette concaténation est nommé en fonction de la date et de l'heure de sa création et en fonction de l'utilisateur de l'application.

Dans l'exemple ci-dessous, la concaténation a été effectuée à 14 h 10 le 7 décembre 2022 par l'utilisateur « LD259969 » à partir des listes consolidées des années 2018 à 2022.

 2022-12-07 1410 Concaténation par LD259969 2018 2019 2020 2021 2022.xlsx

Enfin, tenant compte que les résultats sont écrasés à chaque changement de type de données au lancement de l'application, la liste consolidée et sa décomposition par type de publications est sauvegardée dans un dossier spécifique sous le dossier suivant :



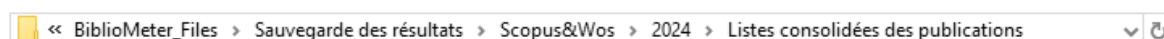
Le nom du dossier spécifique est de la forme :

« <type de données>/<année>/Listes consolidées des publications »

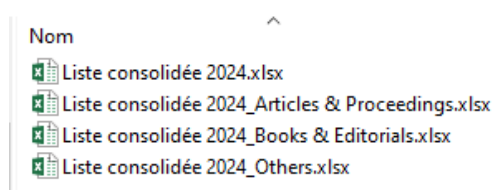
où <année> indique l'année traitée et <type de données> est défini par le type de données sélectionné au lancement de l'application selon :

- « Scopus&Wos » pour la sélection de « Scopus & WoS » ;
- « HalScopus&Wos » pour la sélection de « Scopus-HAL & WoS » ;
- « Wos » pour la sélection de « WoS ».

Par exemple, pour l'analyse conjuguée des extractions de Scopus et WoS et le corpus de l'année 2024, les listes consolidées seront sauvegardées dans le dossier :



Il contient les fichiers suivants :



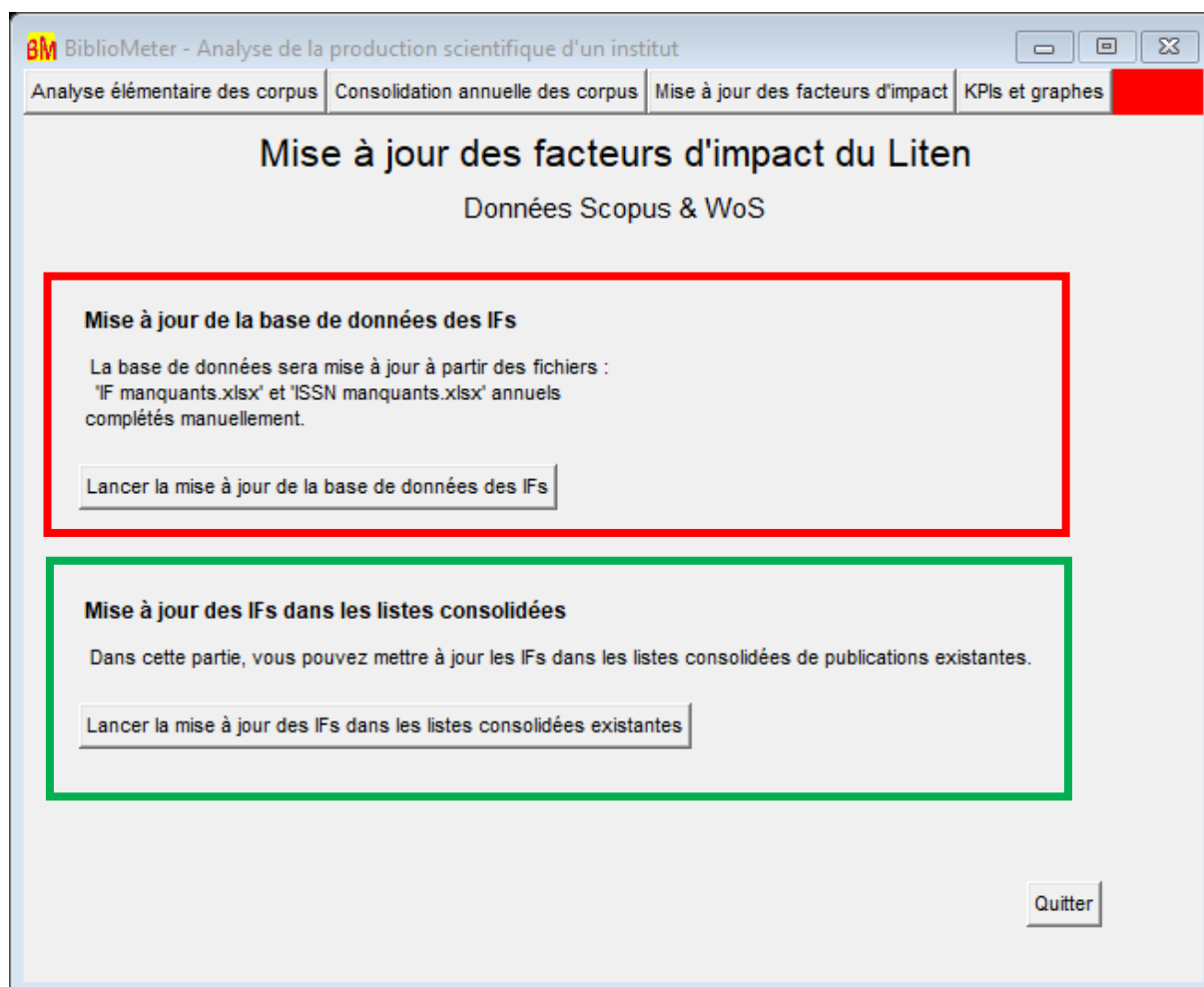
VII. Onglet « Mise à jour des facteurs d'impact »

Cette section de l'application permet de mettre à jour les facteurs d'impact (IF) attribués aux publications au moment de la publication annuelle des facteurs d'impact des journaux par *Clarivate Analytics*.

L'attribution des IFs aux publications au cours de l'année « n » se fait dans les conditions suivantes :

- Les IFs de l'année « n-1 » sont publiés au mois de juillet de l'année « n ».
- Pour les publications de l'année « n-1 » :
 - Jusqu'à juillet de l'année « n », ne peuvent être attribués que les IFs de l'année « n-2 ».
 - À partir de juillet de l'année « n », les IFs attribués peuvent être mis à jour avec les IFs de l'année « n-1 ».
- Pour les publications de l'année « n » :
 - Jusqu'à juillet de l'année « n », ne peuvent être attribués que les IFs de l'année « n-2 ».
 - À partir de juillet de l'année « n », les IFs attribués peuvent être mis à jour avec les IFs de l'année « n-1 »
 - À partir de juillet de l'année « n+1 », les IFs attribués peuvent être mis à jour avec les IFs de l'année « n »

À la sélection de cet onglet, le contenu de la fenêtre est celui montré ci-dessous. La fenêtre comporte 2 actions décrites ci-après.



VII.1. Mise à jour de la base de données des IFs

Le haut de l'onglet (**cadre rouge**) permet d'effectuer la mise à jour de la base de données des IFs en s'appuyant sur les informations apportées aux fichiers « <année>_IF manquants.xlsx » et « <année>_ISSN manquants.xlsx » de toutes les années <année> disponibles dans le dossier « BiblioMeter_files ».

Pour renseigner ces fichiers, vous reporter à l'opération « **Renseignement des fichiers des IFs manquants et des ISSN manquants** » (cf § XV).

Après que ces fichiers sont renseignés, la mise à jour de la base de données des IFs se fait par l'opération suivante :

- Cliquer sur le bouton « **Lancer la mise à jour de la base de données des IFs** ». Cette opération prend quelques secondes et un message informe de sa complétion ou des erreurs rencontrées.

VII.2. Mise à jour des IFs dans les listes consolidées

Le bas de l'onglet (**cadre vert**) sert à mettre à jour les IFs dans les listes consolidées de toutes les années disponibles dans le dossier « BiblioMeter_files » par l'opération suivante :

- Cliquer sur le bouton « **Lancer la mise à jour des IFs dans les listes consolidées existantes** ». Cette opération prend quelques secondes et un message informe de sa complétion ou des erreurs rencontrées.

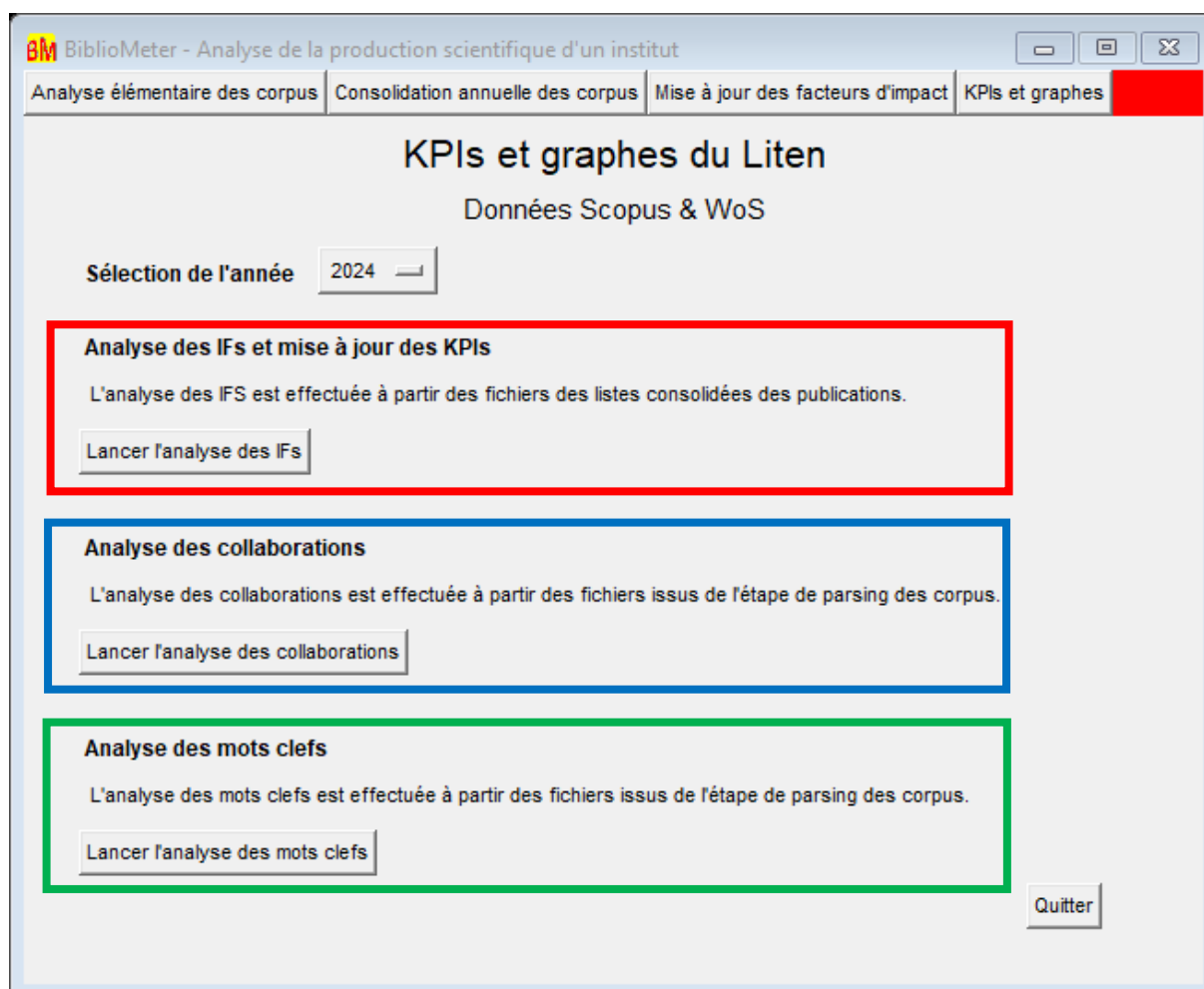
VIII. Onglet « KPIs et graphes »

Cette section de l'application permet de faire l'analyse approfondie de la production scientifique à partir de la liste consolidée des publications. En complément de la liste consolidée, cette analyse peut s'appuyer sur les éléments d'information disponibles dans les dossiers nommés « parsing » et issus de l'étape d'analyse élémentaire des corpus.

Actuellement cette analyse concerne :

- L'analyse des impacts facteurs (IFs) et le calcul d'indicateurs de performance (KPIs) ;
- L'analyse des collaborations du point de vue géographique (couplage par pays et par continent) et la préparation de l'analyse des organismes co-auteurs (normalisation des affiliations) ;
- L'analyse des mots clefs caractérisant les publications de l'Institut.

À la sélection de cet onglet, le contenu de la fenêtre est celui montré ci-dessous. La fenêtre comporte 3 actions décrites ci-après. Ces actions indépendantes sont à effectuer pour l'année sélectionnée par le menu déroulant en face de « **Sélection de l'année** ».



VIII.1. Analyse des IFs et calcul des KPIs

Le haut de l'onglet (**cadre rouge**) permet d'effectuer l'analyse des IFs en s'appuyant sur les informations de la liste consolidée de l'année traitée.

L'analyse des IFs et le calcul des KPIs se fait par l'opération suivante :

- Cliquer sur le bouton « **Lancer l'analyse des IFs** ». Cette opération prend moins d'une minute et un message informe de sa complétion ou des erreurs rencontrées.

À l'issue de cette action, plusieurs fichiers sont créés dans le dossier « **BiblioMeter_Files\<année>\5-Analyses\IFs** » pour chaque département de l'Institut ainsi que des fichiers de synthèse pour l'Institut. Ces fichiers sont de 3 types :

- Fichiers avec extension « **xlsx** » avec 3 colonnes : le nom de journal, le nombre de publications dans le journal, l'IF du journal. L'exemple ci-dessous montre le contenu pour le traitement de l'année 2024 en mars quand les IFs disponibles sont ceux de 2022.

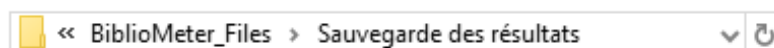
Journal	Number	IF 2022
Physical Chemistry Chemical Physics	1	3,3
Journal of Physical Chemistry C	1	3,7
Journal of The Electrochemical Society	2	3,9

- Fichiers avec l'extension « **html** », bar graphes interactifs correspondant aux mêmes informations que dans les fichiers ci-dessus.
- Fichiers avec l'extension « **png** », images des bar graphes ci-dessus.

De plus, un fichier de synthèse des KPIs est créé ou mis à jour avec les données d'analyse de l'année traitée dans le dossier « **BiblioMeter_Files\BDD multi annuelle** » pour chaque département de l'Institut ainsi que des fichiers de synthèse pour l'Institut. La structure du contenu d'un tel fichier est montrée ci-dessous.

Année de publication	2018	2019	2020	2021	2022	2023	2024
Nombre de publications							
Ouvrages							
Chapitres							
Moyenne de chapitres par ouvrage							
Maximum de chapitres par ouvrage							
Journaux & actes de conférence							
Journaux							
Articles & communications							
Communications							
Communications (%)							
Articles							
Moyenne d'articles par journal							
Maximum d'articles par journal							
Facteur d'impact d'analyse	IF 2018	IF 2019	IF 2020	IF 2021	IF 2022	IF 2022	IF 2022
Facteur d'impact maximum							
Facteur d'impact minimum							
Facteur d'impact moyen							
Articles sans facteur d'impact							
Articles sans facteur d'impact (%)							

Enfin, tenant compte que les résultats sont écrasés à chaque changement de type de données au lancement de l'application, les fichiers avec l'extension « **.xlsx** » sont sauvegardés dans un dossier spécifique sous le dossier suivant :



Pour les fichiers annuels, le nom du dossier spécifique est de la forme :

« <type de données>/<année>/Analyse des facteurs d'impact »

où <année> indique l'année traitée et <type de données> est défini par le type de données sélectionné au lancement de l'application selon :

- « **Scopus&Wos** » pour la sélection de « **Scopus & WoS** » ;
- « **HalScopus&Wos** » pour la sélection de « **Scopus-HAL & WoS** » ;
- « **Wos** » pour la sélection de « **WoS** ».

Par exemple, pour l'analyse conjuguée des extractions de Scopus et WoS et le corpus de l'année 2023, les fichiers seront sauvegardés dans le dossier :

« BiblioMeter_Files » > Sauvegarde des résultats > Scopus&Wos > 2023 > Analyse des facteurs d'impact

La base de données des indicateurs, le nom de dossier spécifique est de la forme :

« <type de données>/Synthèse des indicateurs »

où <type de données> est défini par le type de données sélectionné au lancement de l'application selon :

- « **Scopus&Wos** » pour la sélection de « **Scopus & WoS** » ;
- « **HalScopus&Wos** » pour la sélection de « **Scopus-HAL & WoS** » ;
- « **Wos** » pour la sélection de « **WoS** ».

Par exemple, pour l'analyse conjuguée des extractions de Scopus et WoS, la base de données des indicateurs sera sauvegardée dans le dossier :

« BiblioMeter_Files » > Sauvegarde des résultats > Scopus&Wos > Synthèse des indicateurs

VIII.2. Analyse des collaborations

Le milieu de l'onglet (**cadre bleu**) permet d'effectuer l'analyse des collaborations en s'appuyant sur les informations de la liste consolidée de l'année sélectionnée et sur le fichier « **addresses.dat** », du dossier « **BiblioMeter_Files\<année>\Corpus\deduplication\parsing** » où <année> est l'année traitée.

L'analyse des collaborations se fait par l'opération suivante :

- Cliquer sur le bouton « **Lancer l'analyse des collaborations** ». Cette opération prend plusieurs minutes et un message informe de sa complétion ou des erreurs rencontrées.

À l'issue de cette action, deux dossiers sont créés dans le dossier « **BiblioMeter_Files\<année>\5 - Analyses** » où <année> indique l'année traitée décrits ci-après.

VIII.2.a Analyse géographique

Le dossier qui contient les fichiers relatifs à l'analyse géographique des collaborations est le suivant :

« 5 - Analyses » > Géographique

Il contient les fichiers EXCEL suivants :

- « **Statistiques par continent.xlsx** » avec 3 colonnes : le nom du continent, le nombre de publications dont les auteurs sont issus d'un organisme de ce continent, la liste des identifiants

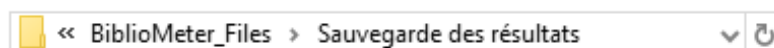
des publications concernées dans la liste consolidées des publications de l'année traitée au format « <aaaa>_### » où <aaaa> indique l'année traitée et « ### » le numéro d'ordre d'extraction de la publication. L'exemple ci-dessous montre le contenu pour le traitement de l'année 2023.

Continent	Nombre de publications	Liste des Pub_ids
Africa	2	2023_055; 2023_065
Asia	14	2023_003; 2023_022; 2023_054; 2023_108; 2023_117; 2023_125; 2023_131; 2023_166; 2023_184; 2023_190; 2023_199; 2023_200; 2023_214; 2023_425
Europe	231	2023_000...2023_425
North-America	6	2023_054; 2023_081; 2023_116; 2023_141; 2023_166; 2023_425
Oceania	1	2023_166
South-America	3	2023_145; 2023_148; 2023_166

- « **Statistiques par pays.xlsx** » avec 3 colonnes : le nom du pays, le nombre de publications dont les auteurs sont issus d'un organisme de ce pays, la liste des identifiants des publications concernées dans la liste consolidées des publications de l'année traitée au format « <aaaa>_### » où <aaaa> indique l'année traitée et « ### » le numéro d'ordre d'extraction de la publication. L'exemple ci-dessous montre le contenu partiel pour le traitement de l'année 2023.

Pays	Nombre de publications	Liste des Pub_ids
Algeria	1	2023_055
Australia	1	2023_166
Austria	2	2023_046; 2023_185
Belgium	7	2023_044; 2023_079; 2023_113; 2023_136; 2023_144; 2023_172; 2023_425
Brazil	1	2023_166
Canada	1	2023_166
Chile	2	2023_145; 2023_148
China	6	2023_022; 2023_117; 2023_125; 2023_166; 2023_190; 2023_425
Denmark	3	2023_117; 2023_125; 2023_214
Finland	2	2023_044; 2023_169
France	231	2023_000...2023_425
Germany	19	2023_041; 2023_044; 2023_054; 2023_058; 2023_085; 2023_107; 2023_121; 2023_133; 2023_136; 2023_138; 2023_144; 2023_145; 2023_156; 2023_166; 2023_169; 2023_175; 2023_190; 2023_214; 2023_425
Greece	1	2023_054
India	1	2023_131
Iran	1	2023_003
Ireland	1	2023_183
Israel	1	2023_214
Italy	16	2023_006; 2023_022; 2023_044; 2023_058; 2023_059; 2023_067; 2023_079; 2023_088; 2023_096; 2023_144; 2023_172; 2023_173; 2023_183; 2023_188; 2023_356; 2023_425
Lithuania	1	2023_044

Enfin, tenant compte que les résultats sont écrasés à chaque changement de type de données au lancement de l'application, ces fichiers sont sauvegardés dans un dossier spécifique sous le dossier suivant :



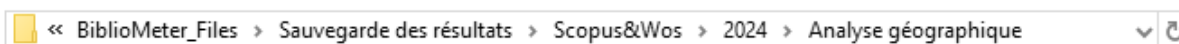
Le nom du dossier spécifique est de la forme :

« <type de données>/<année>/Analyse géographique »

où <année> indique l'année traitée et <type de données> est défini par le type de données sélectionné au lancement de l'application selon :

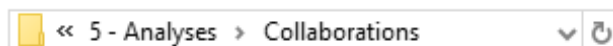
- « **Scopus&Wos** » pour la sélection de « **Scopus & WoS** » ;
- « **HalScopus&Wos** » pour la sélection de « **Scopus-HAL & WoS** » ;
- « **Wos** » pour la sélection de « **WoS** ».

Par exemple, pour l'analyse conjuguée des extractions de Scopus et WoS et le corpus de l'année 2024, les fichiers seront sauvegardés dans le dossier :



VIII.2.b Analyse des collaborations par organisme

Le dossier qui contient les fichiers relatifs à l'analyse des collaborations par organisme est le suivant :



Il contient les fichiers EXCEL destinés à la normalisation des affiliations suivants :

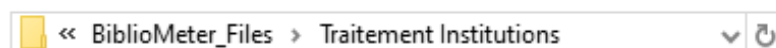
- « **Institutions normalisées.xlsx** » avec 4 colonnes : l'identifiant de la publication dans la liste consolidées des publications de l'année traitée au format « <aaaa>_### » où <aaaa> indique l'année traitée et « ### » le numéro d'ordre d'extraction de la publication, le numéro d'ordre de l'adresse de l'affiliation pour la publication concernée, le nom du pays, la liste des **institutions normalisées** issues de l'analyse de l'adresse de l'affiliation. L'exemple ci-dessous montre le contenu partiel pour le traitement de l'année 2023 ciblant la publication « 2023_000 ».

Pub_id	Idx_address	Country	Institution
2023_000	0	France	Aix Marseille Univ; CNRS Nro; IRD Nro; Collège de France Sch; INRAE Inst; CEREGE Lab
2023_000	1	France	CNRS Nro
2023_000	2	France	Montpellier Univ
2023_000	3	France	UGA Univ; CEA Nro; LITEN Inst

- « **Institutions brutes.xlsx** » avec 4 colonnes : l'identifiant de la publication dans la liste consolidées des publications de l'année traitée au format « <aaaa>_### » où <aaaa> indique l'année traitée et « ### » le numéro d'ordre d'extraction de la publication, le numéro d'ordre de l'adresse de l'affiliation pour la publication concernée, le nom du pays, la liste des **institutions non encore normalisées** issues de l'analyse de l'adresse de l'affiliation. L'exemple ci-dessous montre le contenu partiel pour le traitement de l'année 2023 ciblant la publication « 2023_000 ».

Pub_id	Idx_address	Country	Institution
2023_000	0	France	Empty
2023_000	1	France	CIRAD
2023_000	2	France	Recyclage et Risque; CIRAD
2023_000	3	France	Empty

Dans ce fichier, « **Empty** » dans la colonne « **Institution** » indique que toutes les institutions issues de l'analyse de l'adresse de l'affiliation ont pu être identifiées dans le fichier des institutions normalisées « **<Institut>_Country_affiliations.xlsx** » où <Institut> indique l'Institut sélectionné au lancement de l'application. Ce fichier se trouve dans le dossier suivant :



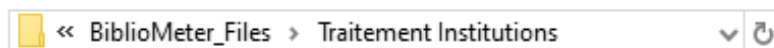
Sinon, la colonne « **Institution** » contient les éléments de l'adresse de l'affiliation qui n'ont pas pu être identifiés dans le fichier des institutions normalisées décrit ci-dessus. Dans ce cas, l'utilisateur doit compléter le fichier « **<Institut>_Country_affiliations.xlsx** » jusqu'à aboutir à une colonne « **Institution** » qui ne contient que la valeur « **Empty** ».

Ce fichier « **<Institut>_Country_affiliations.xlsx** » comporte une feuille par pays et chaque feuille contient les colonnes suivantes : le nom normalisé de l'affiliation dans la colonne « **Norm affiliations** »

puis 1 colonne « **Raw affiliations #** » pour chaque élément d'adresse d'affiliation pouvant être rencontré dans les adresses. L'exemple ci-dessous montre le contenu pour le Danemark :

Norm affiliations	Raw affiliations 1	Raw affiliations 2
AALBORG CSP Firm	AALBORG CSP	
AAU Univ	AAU	Aalborg University
Chem & Biosci Dept	Department Chem & Biosci	Department of Chemistry and Bioscience
Civil Engrn Dept	Department Civil Engrn	Department of Civil Engineering
CNG Dept	CNG	Center for Nanostructured Graphene
Compute Dept	DTU Compute	Department Appl Math & Computer Sci
DTU Univ	DTU	Technical University of Denmark
Electrical Engrn Dept	Department Electrical Engrn	Department of Electrical Engineering
Energy Convers & Storage Dept	Department of Energy Conversion and Storage	Department Energy Convers & Storage
Energy Technol Dept	Department Energy Technol	Department of Energy Technology
EWII Firm	EWII	IRD Fuel Cells
Mads Clausen Inst	Mads Clausen Institute	
Niels Bohr Inst	Niels Bohr Institute	
QDev Dept	Ctr Quantum Devices	Center of Quantum Devices
SDU NanoSYD Dept	SDU NanoSYD	
SDU Univ	SDU	University Southern Denmark
Topsoe Firm	Topsoe	
UCPH Univ	UCPH	University Copenhagen
Vestas Firm	Vestas	

Le nom normalisé des affiliations est composé d'une abréviation choisie par l'utilisateur et d'un type d'institution à définir à partir du fichier « **<Institut>_Institutions_type.xlsx** » où <Institut> indique l'Institut sélectionné au lancement de l'application. Ce fichier se trouve dans le même dossier :



Ce fichier « **<Institut>_Institutions_type.xlsx** » comporte une feuille à 4 colonnes : l'ordre du type d'organisation qui sera utilisé dans la liste des institutions normalisées correspondant à l'adresse concernée, l'abréviation à associer au nom normalisé de l'organisme, le descriptif en anglais du type d'organisation, le descriptif en français du type d'organisation. Ces descriptifs sont donnés pour accompagner le choix de l'utilisateur. L'exemple ci-dessous montre le contenu pour les 7 premiers types d'organisation :

Level	Abbreviation	En significations	Fr significations
1	Gov	Government unit ; Ministry	Ministère ; Centre d'étude et d'expertise
2	Firm	Company; Startup	Inductriel ; Startup
3	Agn	Agency	Agence
4	Univ	University	Université
5	Nro	National Research Organization	Organisation de recherche gouvernementale
6	Iro	International Research Organization	Organisation de recherche internationale
7	Lsf	Large scale facility	Grand instrument

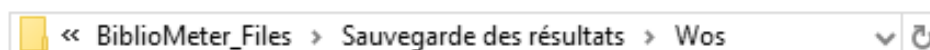
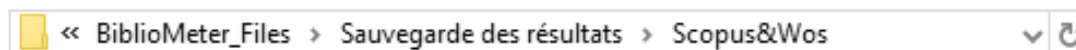
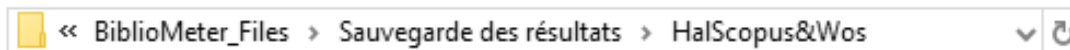
Pour tenir compte des compléments apportés au fichier « **<Institut>_Country_affiliations.xlsx** », il est nécessaire de **relancer l'analyse des collaborations** pour l'année traitée. Attention, ces opérations sont relativement consommatrices de temps.

L'exploitation du fichier « **Institutions normalisées.xlsx** » pour une analyse des collaborations reste à définir après une préparation plus avancée de ces fichiers.

Les fichiers du dossier résultant de l'analyse des collaborations par organisme ne sont conservés que manuellement pour le type de données « Scopus-Hal & WoS » sélectionné au lancement de l'application en copiant le dossier de l'année traitée suivant :



dans les dossiers de l'année traitée des dossiers suivants :



VIII.3. Analyse des mots clefs

Le bas de l'onglet (**cadre vert**) permet d'effectuer l'analyse des mots clefs caractérisant les publications de l'Institut en s'appuyant sur les informations de la liste consolidée de l'année sélectionnée et sur les fichiers « **authorskeywords.dat** », « **journalkeywords.dat** » et « **titlekeywords.dat** » du dossier « **BiblioMeter_Files\<année>\Corpus\deduplication\parsing** » où <année> est l'année traitée.

L'analyse des mots clefs se fait par l'opération suivante :

- Cliquer sur le bouton « **Lancer l'analyse des mots clefs** ». Cette opération prend moins d'une minute et un message informe de sa complétion ou des erreurs rencontrées.

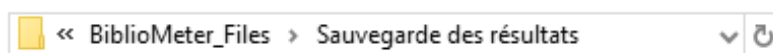
À l'issue de cette action, plusieurs fichiers sont créés dans le dossier « **BiblioMeter_Files\<année>5-Analyses\Mots clefs** » pour chaque département de l'Institut ainsi que des fichiers de synthèse pour l'Institut. Ces fichiers sont de 2 types :

- Fichiers avec extension « **xlsx** » avec 2 colonnes : le mot clef (« **Keyword** »), son niveau d'occurrence exprimé en nombre d'articles (« **Weight** »). L'exemple ci-dessous montre le contenu type. Le nom des fichiers rappelle le département, l'année du corpus et le type de mots clefs avec « **AK** », les mots clefs définis par les auteurs, « **IK** », les mots clefs attribués par la base de données (mots clefs indexés) et « **TK** », les mots clefs extraits des titres des publications.

Keyword	Weight
approach	1
cell	1
diffusion	3
electrolyte	3
layer	2
material	1
mechanical	3
modelling	1
solid	2
temperature	1
thermal	1

- Fichiers avec extension « **png** », images des représentations graphiques en nuages de mots correspondants aux fichiers ci-dessus.

Enfin, tenant compte que les résultats sont écrasés à chaque changement de type de données au lancement de l'application, les fichiers avec l'extension « **xlsx** » sont sauvegardés dans un dossier spécifique sous le dossier suivant :



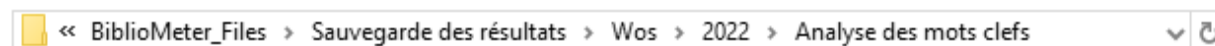
Le nom du dossier spécifique est de la forme :

« **<type de données>/<année>/Analyse des mots clefs** »

où <année> indique l'année traitée et <type de données> est défini par le type de données sélectionné au lancement de l'application selon :

- « **Scopus&Wos** » pour la sélection de « **Scopus & WoS** » ;
- « **HalScopus&Wos** » pour la sélection de « **Scopus-HAL & WoS** » ;
- « **Wos** » pour la sélection de « **WoS** ».

Par exemple, pour l'analyse des extractions de WoS seul et le corpus de l'année 2022, les fichiers seront sauvegardés dans le dossier :



IX. Attribution des OTPs

Ce mode opératoire explique comment sélectionner un OTP pour les publications scientifiques de l'Institut sur un document Excel généré par l'utilisation de l'application « **BiblioMeter** ». Si le fichier n'a pas encore été généré, veuillez-vous rendre au **paragraphe VI.3** pour sa création.

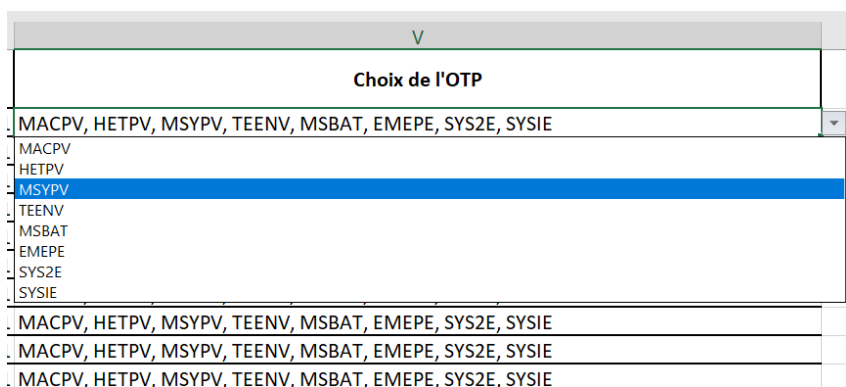
1. Ouvrir le dossier « **aaaa** » de l'année des publications auxquelles vous souhaitez rajouter les OTP dans le dossier de travail « **BiblioMeter_Files** ».

2. Ouvrir le dossier « **2 - OTP** »

3. Ouvrir le fichier de votre département parmi les fichiers nommés « **fichier_ajout_OTP_<dept>.xlsx** » où <dept> désigne le nom du département.

4. Ajouter les OTP :

Chacun de ces quatre fichiers Excel possède les informations nécessaires à l'identification de la publication et de ses auteurs. La colonne qui nous intéresse plus particulièrement ici est celle intitulée « **Choix de l'OTP** ». En cliquant sur la cellule de la publication en question, il est possible de faire apparaître un menu déroulant qui propose l'ajout d'un OTP à la cellule.



Le menu déroulant n'est qu'une aide qui vise à faciliter le choix de l'OTP, vous pouvez écrire ce que vous voulez à la place si les propositions faites ne vous conviennent pas.

Le menu déroulant vous propose également la valeur « **Invalide** ». Cette valeur permet d'identifier les publications à ne pas prendre en compte lors de la consolidation de la liste des publications.

Une fois cette sélection faite pour toutes les publications, la colonne « **Choix de l'OTP** » ne devrait plus contenir qu'une valeur du menu déroulant comme dans l'exemple suivant :

V
Choix de l'OTP
MACPV
HETPV
MSYPV
TEENV
MSBAT
MACPV
HETPV
MSYPV
TEENV
MSBAT
HETPV
MSYPV
TEENV
MSBAT

5. Sauvegarder vos modifications :

Sauvegarder le document **en rajoutant « _ok »** à son nom et le fermer si vous êtes satisfait des choix des OTPs que vous avez fait.

Ex : « **fichier_ajout_OTP_DTS.xlsx** » devient « **fichier_ajout_OTP_DTS_ok.xlsx** »

Si la sélection de la valeur d'OTP n'est **effectuée que pour une partie des publications**,
sauvegarder le document sans changer son nom.

Remarques :

- La création des fichiers pour l'attribution des OTPs est nécessaire à la poursuite du traitement mais l'attribution des OTPs elle-même est facultative. Si les fichiers avec l'ajout de « **_ok** » ne sont pas disponibles, ce sont les fichiers initiaux qui seront utilisés dans la suite du traitement.
- Les OTPs attribués sont conservés dans le fichier « **OTPs conservés** » sauvegardé dans le dossier « **BiblioMeter_Files\<année>\4 – Informations** », où <année> est l'année sélectionnée, en correspondance avec l'identifiant universel de l'article et seront automatiquement attribués lors d'un nouveau traitement du corpus de l'année, y compris la valeur « **Invalidé** » pour les publications à supprimer.

Retour : VI.3

X. Extraction d'un corpus de l'année

Cette section décrit les modes opératoires pour l'extraction des données bibliographiques des bases Scopus et Web of Science. Cette extraction concerne l'année à analyser et peut se faire à tout moment de l'année à une fréquence définie par l'Institut pour le suivi de sa production scientifique.

X.1. Extraction pour l'Institut LITEN

Cette section décrit les modes opératoires pour l'extraction des données bibliographiques des bases Scopus et Web of Science pour le LITEN. **La fréquence d'extraction choisie est mensuelle** au cours de l'année d'analyse.

X.1.a Extraction depuis Scopus

1. Recherche sous Scopus via le lien url (connexion CEA requise)
2. Cliquer sur l'onglet « Advanced Search » et rentrer la requête suivante: AFFILORG(LITEN OR INES) AND AFFILCOUNTRY(FRANCE) AND PUBYEAR IS 20xx <i>Rem : cela ne suffit pas à éliminer les publications d'autres organismes pouvant avoir un sigle « INES » (Espagne et Allemagne) « Liten » (Maroc). Une recherche manuelle ultérieure sera nécessaire pour éliminer ces cas</i>
3. Lancer la recherche
4. Sélectionner « All »
5. Sélectionner « Export »
6. Choisir le format « csv »
7. Dans les informations à exporter tout sélectionner sauf (pour ne pas alourdir le fichier importé, sans conséquences pour l'application elle-même) : <ul style="list-style-type: none">• Tout « Funding details »• Dans « Other information » :<ul style="list-style-type: none">○ « Tradenames & manufacturers »○ « Accession numbers & chemicals »○ « Conference information »
8. Transférer le fichier « .csv » des « Téléchargements » vers le répertoire « BiblioMeter_Files/Extractions Institut/ScopusExtractions_Files/20xx »
9. Renommer le fichier « aaaa-mm-jj scopus.csv » où <aaaa-mm-jj> correspond à la date d'extraction
10. Déplacer le fichier de l'extraction précédente dans le répertoire d'archives « BiblioMeter_Files/Extractions Institut/ScopusExtractions_Files/20xx/Archives »

X.1.b Extraction depuis WoS

1. Recherche sous WoS via le lien url
2. Cliquer sur l'onglet « Advanced Search » et rentrer la requête suivante : (AD=(CEA, LITEN) OR AD=(INES)) AND AD=(FRANCE) AND PY=20xx <i>Rem : cela ne suffit pas à éliminer les publications d'autres organismes pouvant avoir un sigle « INES » (Espagne et Allemagne) « Liten » (Maroc). L'application éliminera ces publications.</i>
3. Cliquer sur Search (puis éventuellement sur le nombre de résultats pour visualiser les publications et vérifier rapidement que la requête est correcte)
4. Cliquer sur « Export »
5. Sélectionner dans le menu déroulant : « Tab Delimited File »
6. Dans la nouvelle fenêtre : a. Sélectionner « Records from xxxx to xxxx » (max 500 par export) b. Sélectionner dans menu déroulant « Record content » : « Full record and Cited References » c. Click sur « Export »
7. Transférer le fichier « .txt » des « Téléchargements » vers le répertoire « BiblioMeter_Files/Extractions Institut/WosExtractions_Files/20xx »
8. Renommer le fichier « aaaa-mm-jj wos.txt » où <aaaa-mm-jj> correspond à la date d'extraction
9. Déplacer le fichier de l'extraction précédente dans le répertoire d'archives « BiblioMeter_Files/Extractions Institut/WosExtractions_Files/20xx/Archives »

Signification des colonnes dans WoS

Liste indicative, peut ne pas être conservée d'une année sur l'autre : vérifier !

TS = Topic	SO = Publication Name	SA = Street Address	SU = Research Area
TI = Title	DO = DOI	CI = City	WC = Web of Science Category
AB = Abstract	PY = Year Published	PS = Province/State	IS = ISSN/ISBN
AU = Authors	CF = Conference	CU = Country/Region	UT = Accession Number
AI = Author Identifiers	AD = Address	ZP = Zip/Postal Code	PMID = PubMed ID
AK = Author Keywords	OG = Affiliations	FO = Funding Agency	DOP = Publication Date
GP = Group Author	OO = Organization	FG = Grant Number	PUBL = Publisher
ED = Editor	SG = Suborganization	FD = Funding Details	ALL = All Fields
KP = Keyw ord Plus		FT = Funding Text	FPY = Final publication year

Retour : V.1

X.2. Extraction pour l'Institut LETI

Cette section décrit les modes opératoires pour l'extraction des données bibliographiques des bases Scopus et Web of Science pour le LETI.

X.2.a Extraction depuis Scopus

1. Recherche sous Scopus via le lien url (connexion CEA requise)
2. Rentrer directement dans « Advanced search » : AF-ID ("CEA LETI " 60007823) OR AFFIL ((leti OR ("Electronic Laboratory of Information Technology") OR ("Laboratoire d'Electronique et de Technologie de l'Information")) AND france) AND PUBYEAR = 20xx <i>Rem : cela ne suffit pas à éliminer les publications d'autres organismes pouvant avoir un sigle avoir un sigle « LETI » (Tunisie/Sfax et Russie/St Petersburg). A priori, l'application éliminera ces publications - Requête nommée Q11</i>
3. Lancer la recherche
4. Sélectionner « All »
5. Sélectionner « Export »
6. Choisir le format « csv »
7. Dans les informations à exporter tout sélectionner sauf (pour ne pas alourdir le fichier importé, sans conséquences pour l'application elle-même) : <ul style="list-style-type: none">• Tout « Funding details »• Dans « Other information » :<ul style="list-style-type: none">○ « Tradenames & manufacturers »○ « Accession numbers & chemicals »○ « Conference information »
8. Transférer le fichier « .csv » des « Téléchargements » vers le répertoire « BiblioMeter_Files/Extractions Institut/ScopusExtractions_Files/20xx »
9. Renommer le fichier « aaaa-mm-jj scopus.csv » où <aaaa-mm-jj> correspond à la date d'extraction
10. Déplacer le fichier de l'extraction précédente dans le répertoire d'archives « BiblioMeter_Files/Extractions Institut/ScopusExtractions_Files/20xx/Archives »

X.2.b Extraction depuis WoS

1. Recherche sous WoS via le lien url
2. Cliquer sur l'onglet « Advanced Search » et rentrer la requête suivante : AD=(LETI) AND AD=(FRANCE) AND PY=20xx <i>Rem : cela ne suffit pas à éliminer les publications d'autres organismes pouvant avoir un sigle « LETI » (Tunisie/Sfax et Russie/St Petersburg). L'application éliminera ces publications - Requête nommée Q4</i>
3. Cliquer sur Search (puis éventuellement sur le nombre de résultats pour visualiser les publications et vérifier rapidement que la requête est correcte)
4. Cliquer sur « Export »
5. Sélectionner dans le menu déroulant : « Tab Delimited File »
6. Dans la nouvelle fenêtre : a. Sélectionner « Records from xxxx to xxxx » (max 500 par export) b. Sélectionner dans menu déroulant « Record content » : « Full record and Cited References » c. Cliquez sur « Export »
7. Transférer le fichier « .txt » des « Téléchargements » vers le répertoire « BiblioMeter_Files/Extractions Institut/WosExtractions_Files/20xx »
8. Renommer le fichier « aaaa-mm-jj wos.txt » où <aaaa-mm-jj> correspond à la date d'extraction
9. Déplacer le fichier de l'extraction précédente dans le répertoire d'archives « BiblioMeter_Files/Extractions Institut/WosExtractions_Files/20xx/Archives »

Signification des colonnes dans WoS

Liste indicative, peut ne pas être conservée d'une année sur l'autre : vérifier !

TS = Topic	SO = Publication Name	SA = Street Address	SU = Research Area
TI = Title	DO = DOI	CI = City	WC = Web of Science Category
AB = Abstract	PY = Year Published	PS = Province/State	IS = ISSN/ISBN
AU = Authors	CF = Conference	CU = Country/Region	UT = Accession Number
AI = Author Identifiers	AD = Address	ZP = Zip/Postal Code	PMID = PubMed ID
AK = Author Keywords	OG = Affiliations	FO = Funding Agency	DOP = Publication Date
GP = Group Author	OO = Organization	FG = Grant Number	PUBL = Publisher
ED = Editor	SG = Suborganization	FD = Funding Details	ALL = All Fields
KP = Keyw ord Plus		FT = Funding Text	FPY = Final publication year

Retour : V.1

XI. Mise à jour du corpus de l'année

Cette section décrit le mode opératoire pour la mise à jour des extractions des données bibliographiques des bases Scopus et Web of Science. Cette mise à jour est nécessaire pour tenir compte des publications qui ne sont répertoriées que plusieurs mois après la fin de l'année d'analyse. C'est le cas de certaines publications liées à des conférences ayant eu lieu en fin d'année.

Pour effectuer cette mise à jour pour Scopus, les étapes sont les suivantes :

1. Suivre le mode opératoire d'extraction depuis Scopus pour l'Institut jusqu'à l'**étape 8** ;
2. Renommer le fichier « **aaaa-final scopus.csv** » où <aaaa> correspond à l'année d'analyse ;
3. Déplacer le fichier de l'extraction précédente dans le répertoire d'archives « **BiblioMeter_Files/Extractions Institut/ScopusExtractions_Files/20xx/Archives** ».

Pour effectuer cette mise à jour pour WoS, les étapes sont les suivantes :

1. Suivre le mode opératoire d'extraction depuis WoS pour l'Institut jusqu'à l'**étape 7** ;
2. Renommer le fichier « **aaaa-final wos.txt** » où <aaaa> correspond à l'année d'analyse ;
3. Déplacer le fichier de l'extraction précédente dans le répertoire d'archives « **BiblioMeter_Files/Extractions Institut/WosExtractions_Files/20xx/Archives** ».

XII. Complément HAL du corpus de l'année

En cours de rédaction

XIII. Mise à jour du fichier des effectifs de l'Institut

Chaque trimestre le service RH envoie à la personne en charge (Direction scientifique) de la mise à jour des effectifs de l'Institut, un document qui inclura les feuilles mensuelles des effectifs pour les mois à mettre à jour.

Exemple :

- En Janvier 2022 → Octobre, Novembre et Décembre 2021
- En Avril 2022 → Janvier, Février et Mars 2022
- Et ainsi de suite ...

- | |
|--|
| <ol style="list-style-type: none">1. Ouvrir le fichier envoyé par le service RH et vérifier que :<ul style="list-style-type: none">• Le nom des feuilles soit au format « mmaaaa » comme dans les exemples suivants :
<i>Janvier 2022 → 012022, Avril 2021 → 042021, Décembre 2023 → 122023 ...</i> <p>Si ce n'est pas le cas, il faut :</p> <ul style="list-style-type: none">- Changer le nom des feuilles en suivant le modèle présenté ci-dessus.- Sauvegarder le document.
<ul style="list-style-type: none">• La liste des informations contenues dans ces feuilles doit être au minimum celles indiquées en gras parmi les suivantes et avec exactement le même intitulé (y compris les espaces) : |
|--|

	<table border="1"> <tr><td>Matricule</td></tr> <tr><td>Nom</td></tr> <tr><td>Prénom</td></tr> <tr><td>Sexe(lib)</td></tr> <tr><td>Nationalité (lib)</td></tr> <tr><td>Catégorie de salarié (lib)</td></tr> <tr><td>Statut de salarié (lib)</td></tr> <tr><td>Filière classement (lib)</td></tr> <tr><td>Qualification classement (lib)</td></tr> <tr><td>Spécialité poste (lib)</td></tr> <tr><td>Nature de contrat (lib)</td></tr> <tr><td>Annexe classement</td></tr> <tr><td>Date d'effet classement</td></tr> <tr><td>Date début contrat</td></tr> <tr><td>Date dernière entrée</td></tr> <tr><td>Date de fin de contrat</td></tr> </table>	Matricule	Nom	Prénom	Sexe(lib)	Nationalité (lib)	Catégorie de salarié (lib)	Statut de salarié (lib)	Filière classement (lib)	Qualification classement (lib)	Spécialité poste (lib)	Nature de contrat (lib)	Annexe classement	Date d'effet classement	Date début contrat	Date dernière entrée	Date de fin de contrat		<table border="1"> <tr><td>Dpt/DOB (lib court)</td></tr> <tr><td>Service (lib court)</td></tr> <tr><td>Laboratoire (lib court)</td></tr> <tr><td>Laboratoire (lib long)</td></tr> <tr><td>N° id du poste budgétaire</td></tr> <tr><td>Unité structurelle</td></tr> <tr><td>Unité structurelle (code R3)</td></tr> <tr><td>Nature de dépenses</td></tr> <tr><td>Taux d'occupation de poste</td></tr> <tr><td>TA</td></tr> <tr><td>Taux d'Activité</td></tr> <tr><td>Règle de plan de roulement (lib)</td></tr> <tr><td>Regpt PR niveau 1(lib)</td></tr> <tr><td>Date de naissance</td></tr> <tr><td>Age</td></tr> <tr><td>Tranche d'age (5 ans)</td></tr> </table>	Dpt/DOB (lib court)	Service (lib court)	Laboratoire (lib court)	Laboratoire (lib long)	N° id du poste budgétaire	Unité structurelle	Unité structurelle (code R3)	Nature de dépenses	Taux d'occupation de poste	TA	Taux d'Activité	Règle de plan de roulement (lib)	Regpt PR niveau 1(lib)	Date de naissance	Age	Tranche d'age (5 ans)	
Matricule																																				
Nom																																				
Prénom																																				
Sexe(lib)																																				
Nationalité (lib)																																				
Catégorie de salarié (lib)																																				
Statut de salarié (lib)																																				
Filière classement (lib)																																				
Qualification classement (lib)																																				
Spécialité poste (lib)																																				
Nature de contrat (lib)																																				
Annexe classement																																				
Date d'effet classement																																				
Date début contrat																																				
Date dernière entrée																																				
Date de fin de contrat																																				
Dpt/DOB (lib court)																																				
Service (lib court)																																				
Laboratoire (lib court)																																				
Laboratoire (lib long)																																				
N° id du poste budgétaire																																				
Unité structurelle																																				
Unité structurelle (code R3)																																				
Nature de dépenses																																				
Taux d'occupation de poste																																				
TA																																				
Taux d'Activité																																				
Règle de plan de roulement (lib)																																				
Regpt PR niveau 1(lib)																																				
Date de naissance																																				
Age																																				
Tranche d'age (5 ans)																																				
<p>Si ce n'est pas le cas, il faut :</p> <ul style="list-style-type: none"> - Changer le nom des colonne en recopiant ceux donnés ci-dessus. - Sauvegarder le document. 																																				
<p>2. Ouvrir la zone de dépôt du fichier dans « BiblioMeter_Files » suivante :</p> <ul style="list-style-type: none"> - « Parametres Institut\ Effectifs de consolidation » 																																				
<p>3. Supprimer l'ancien document présent dans ce dossier,</p> <p>!!! ATTENTION !!! SEULEMENT LE NOUVEAU DOCUMENT DOIT ÊTRE DANS LE DOSSIER.</p>																																				
<p>4. Déposer le nouveau fichier dans ce dossier</p>																																				
<p>5. La mise à jour des effectifs sera finalisée par l'application BiblioMeter, lors de la procédure de croisement entre la liste des publications et les effectifs de l'Institut.</p>																																				

Retour : VI.1

XIV. Réduction du fichier « orphan.xlsx »

Après le croisement de la liste des publications avec les effectifs de l'Institut, le fichier « **orphan.xlsx** » contient les auteurs affiliés à l'Institut mais absents du fichier des effectifs. Différentes raisons peuvent expliquer cette absence. Il s'agit d'identifier **la raison pour chacun des auteurs et chacune des publications**. Puis, les corrections à apporter sont indiquées dans l'un des 3 fichiers du dossier « \BiblioMeter_Files\Traitement Orphan » qui se nomment :

- « **Autres corrections.xlsx** » ;
- « **Effectifs additionnels.xlsx** » ;
- « **Orthographe.xlsx** ».

La première étape consiste à :

1. **Ouvrir le fichier « orphan.xlsx »** du dossier « \BiblioMeter_Files\<année>\0- BDD multi mensuelle » ;
2. **Utiliser le DOI** de la publication pour y accéder sur le site de l'éditeur ;
3. **Vérifier les auteurs et leurs affiliations.**

Plusieurs cas peuvent se présenter et les opérations correspondantes sont décrites ci-après.

XIV.1. La liste des auteurs et leurs affiliations sont ambiguës

Par exemple, les affiliations sont listées dans un encadré en bas de première page pour certains éditeurs. Un autre exemple est l'inversion des noms et prénoms. Cela peut-être à l'origine d'erreurs dans les bases de données Scopus et WoS.

Pour prendre en compte ces erreurs, il convient d'identifier la correction à apporter et de l'indiquer dans la feuille nommée « **Spécifique par publi** » du fichier « **Autres corrections.xlsx** ». Les informations à apportées sont les suivantes :

Colonne	Information	Localisation
Nom pub	Nom de l'auteur incorrect	dans « orphan.xlsx »
Initiales pub	Initiales de l'auteur incorrect	
Matricule	Matricule de l'auteur réel	dans « All_effectifs.xlsx »
Nom eff	Nom de l'auteur réel	
Prénom eff	Prénom de l'auteur réel	
Initiales eff	Initiales de l'auteur réel	
Dept	Département de l'auteur réel	
Pub_id	Pub_id de la publication	dans « orphan.xlsx »
Année pub	Année de la publication	
Hash_id	Hash_id de la publication	dans « hash_id.xlsx »
Commentaire	Type d'erreur	<i>Facultatif</i>

Rem : la feuille comporte des corrections déjà indiquées qui peuvent servir d'exemple.

Ces corrections ne sont valables que pour une publication donnée identifiée par son « **Hash_id** ». Pour **trouver le « Hash_id »** d'une publication :

1. Copiez le « **Pub_id** » de la publication donné dans le fichier « **orphan.xlsx** » ;
2. Ouvrez le fichier « **hash_id.xlsx** » du dossier « **\BiblioMeter_Files\<année>\0- BDD multi mensuelle** » ;
3. Copiez le « **Hash_id** » correspondant au « **Pub_id** » dans ce fichier.

Cette identification permet d'effectuer la correction des auteurs quel que soit la version de l'extraction des bases de données Scopus et WoS (le « Pub-id » n'est pas forcément conservé d'une extraction à l'autre).

XIV.2. Les auteurs recherchés ont plusieurs affiliations dont celle de l'Institut

Ceux sont donc des collaborateurs externes, en général du milieu académique (CNRS, Universités...) mais parfois de partenaires industriels, et qui ne sont donc pas dans le fichier des effectifs de l'Institut. Il s'agit de ne pas les retenir comme auteurs de l'Institut sauf pour le cas particulier où ces personnes ont une adresse « outlook » à l'Institut comme par exemple les doctorants sous contrat extérieur et pour le LETI, les collaborateurs du LTM, de ST Microelectronics... Ces derniers cas sont traités dans le paragraphe **XIV.3** ci-après.

Pour prendre en compte les personnes à exclure, il convient d'indiquer ces auteurs dans la feuille nommée « **Externes LITEN** » du fichier « **Autres corrections.xlsx** ». Les informations à apportées sont les suivantes :

Colonne	Information	Localisation
Nom pub	Nom de l'auteur à retirer	dans « orphan.xlsx »
Initiales pub	Initiales de l'auteur à retirer	
Affiliation	Sigle du laboratoire ou de l'entreprise	<i>Facultatif</i>

Ces auteurs seront retirés de la liste des auteurs de l'Institut à chaque fois qu'ils apparaissent dans les liste de publications quel que soit l'année considérée.

XIV.3. Les auteurs recherchés n'ont qu'une seule affiliation, celle de l'Institut

Soit l'auteur est sous contrat externe et ne figure donc pas dans le fichier des effectifs, soit il y a une différence d'orthographe entre l'identification dans la publication et l'identification dans le fichier des effectifs.

- Auteur sous contrat externe

C'est le cas des collaborateurs qui sont sur site et qui ont une adresse « outlook » comme par exemple pour le LETI, les collaborateurs du LTM, de ST Microelectronics et de CLINATEC.

C'est aussi le cas des intérimaires que l'on ne peut identifier qu'au niveau des unités d'accueil en contactant directement les co-auteurs LITEN quand ils existent, sinon l'assistante de l'unité (mais ce cas est rarement rencontré).

Pour prendre en compte ce type de personnel, il convient d'indiquer ces auteurs dans la feuille nommée « **Autres externes** » du fichier « **Effectifs additionnels.xlsx** ». Les informations à apportées

sont les mêmes que dans le tableau donné plus loin en précisant en plus le statut du salarié dans la colonne « Statut ».

C'est surtout le cas des doctorants de type « ADEME », « CIFRE » ou « BDI CNRS » par exemple. Pour le LITEN, ils sont identifiés dans la base de données des doctorants maintenue à jour par la Direction Scientifique de l'Institut.

Pour prendre en compte les doctorants externes, il convient d'indiquer ces auteurs dans la feuille nommée « **Doctorants externes** » du fichier « **Effectifs additionnels.xlsx** ». Les informations à apportées sont les suivantes :

Colonne	Information	Localisation
Nom pub	Nom de l'auteur	dans « orphan.xlsx »
Initiales pub	Initiales de l'auteur	
Nom eff	Nom du doctorant	Dans la base de données des doctorants
Initiales eff	Initiales du doctorant	
Prénom eff	Prénom du doctorant	
Laboratoire	Dept/Serv/Labo	
Promotion	Année de promotion	
Année fin de contrat	Année fin de contrat	
Date début contrat	Date début contrat	
Date de fin de contrat	Date de fin de contrat	
Sexe	Sexe	<i>Facultatif</i>
Nationalité	Nationalité	<i>Facultatif</i>

Rem : Les colonnes grisées sont renseignées automatiquement.

Plusieurs orthographes d'identification peuvent être rencontrées dans les publications pour un même doctorant. Dans ce cas, **chaque orthographe donnera lieu à une ligne se différenciant des autres uniquement par le contenu des 2 premières colonnes** de la feuille à renseigner.

- Différence d'orthographe

Une différence d'orthographe entre l'identification de l'auteur dans la publication et l'identification du salarié dans le fichier des effectifs demande une recherche intégrant différents types de différences :

- Nom composé/nom simple ;
- Prénom composé/ prénom simple ;
- Nom composé et prénom simple/Nom simple et prénom composé
- Présence/absence de tiré de liaison ;
- Présence/absence d'accentuation
- ...

Si le prénom est connu, commencer par sélectionner tous les salariés de ce prénom peut simplifier la recherche. D'autre part, la recherche doit partir de la feuille de l'année de publication et remonter dans le temps jusqu'à une profondeur suffisante (jusqu'à 10 ans dans certains cas).

Pour prendre en compte ces différences d'orthographe, il convient d'indiquer ces auteurs dans le fichier « **Orthographe.xlsx** ». Les informations à apportées sont les suivantes :

Colonne	Information	Localisation
Nom pub	Nom de l'auteur incorrect	dans « orphan.xlsx »
Initiales pub	Initiales de l'auteur incorrect	
Matricule	Matricule du salarié	dans « All_effectifs.xlsx »
Nom eff	Nom du salarié	
Prénom eff	Prénom du salarié	
Initiales eff	Initiales du salarié	

Retour : VI

XV. Renseignement des fichiers des IFs manquants et des ISSN manquants

Lorsque la liste consolidée des publications d'une année est construite, les IFs des journaux disponibles dans la base de données des IFs sont attribués aux publications de ces journaux. Il s'agit d'une part, de l'IF de l'année de publication et d'autre part, de l'IF de la dernière année disponible dénommé « IF de l'année en cours ».

Pour chaque année **<année>** :

- **L'IF de l'année de publication** est normalement disponible dans la base de données des IFs dans la feuille **<année>** à moins qu'une nouvelle extraction des bases de données Scopus ou WoS n'ait ajouté une publication dans un journal qui n'était pas présent dans la liste précédente. Dans ce cas, le journal est présent dans l'un des 2 fichiers « **<année>_IF manquants.xlsx** » et « **<année>_ISSN manquants.xlsx** » selon que son ISSN indiqué par Scopus ou WoS est trouvé dans la base de donnée des IFs ou pas.
- **L'IF de l'année en cours** n'est peut-être pas disponible dans la base de données des IFs à la feuille de la dernière année d'IFs disponible. Le journal est présent dans l'un des 2 fichiers « **<année>_IF manquants.xlsx** » et « **<année>_ISSN manquants.xlsx** » selon que son ISSN indiqué par Scopus ou WoS est trouvé dans la base de donnée des IFs ou pas.

Le renseignement du fichier « **<année>_IF manquants.xlsx** » consiste à :

1. Ouvrir le fichier « **<année>_IF manquants.xlsx** » du dossier « **\BiblioMeter_Files\<année>\3- Résultats Finaux** » ;
2. Utiliser l'ISSN ou l'e-ISSN du journal pour trouver ses IFs sur le site de « [Clarivate Analytics](#) » pour l'année de publication et pour l'année en cours ;
3. Reporter les valeurs trouvées dans les colonnes ad-hoc du fichier « **<année>_IF manquants.xlsx** » ;
4. Remplacer « **unknown** » par « **Not available** », quand ces valeurs d'ISSN ou d'e-ISSN n'existent pas ;
5. Sauvegarder le fichier sous le même nom.

Le renseignement du fichier « **<année>_ISSN manquants.xlsx** » consiste à :

1. Ouvrir le fichier « **<année>_ISSN manquants.xlsx** » du même dossier « **\BiblioMeter_Files\<année>\3- Résultats Finaux** » ;
2. Utiliser l'ISSN du journal issu de la source pour identifier s'il s'agit de son ISSN ou de son e-ISSN sur le site « ISSN portal » ;
3. Déterminer si l'ISSN ou l'e-ISSN sont disponibles ou non sur le site « ISSN portal » ;
4. Renseigner les colonnes ISSN et e-ISSN en conformité avec cette recherche en remplaçant « **unknown** » par « **Not available** » si l'un ou l'autre des identifiants n'existe pas (ex : ISSN inexistant pour journaux uniquement en ligne) ;
5. Utiliser l'ISSN ou l'e-ISSN du journal pour trouver ses IFs sur le site de « [Clarivate Analytics](#) » pour l'année de publication et pour l'année en cours ;
6. Reporter les valeurs trouvées dans les colonnes ad-hoc du fichier « **<année>_IF manquants.xlsx** » ;
7. Remplacer « **unknown** » par « **Not available** », quand ces valeurs d'IF n'existent pas ;
8. Sauvegarder le fichier sous le même nom.

Une fois ces fichiers complétés, il convient d'utiliser l'onglet « Mise à jour des facteurs d'impact » de l'application (cf § VII) pour mettre à jour la base de données des IFs avec le contenu de ces fichiers, puis de mettre à jour toutes les listes consolidées des publications disponibles avec les IFs complétés. Après cette opération, les nouveaux fichiers « <année>_IF manquants.xlsx » et « <année>_ISSN manquants.xlsx » existent mais sont vides.

Une mention « **Not analysed** » est automatiquement affectée aux documents du type « Conference paper », « Book », « Book chapter » et Co. qui ne sont pas pris en compte dans l'analyse des IFs.

XVI. Préparation annuelle de la base de données des IFs

Chaque année n , au mois de juillet, les **IFs des journaux calculés pour l'année $n-1$** sont disponibles sur le site de « Clarivate Analytics ». La préparation de la base de données consiste à créer la feuille correspondant à l'année $n-1$ dans le fichier Excel qui contient la base de données des IFs.

Pour procéder à cette préparation **en juillet de l'année n** , effectuez les opérations suivantes :

1. Ouvrez le document Excel « **<institut>_IF all years.xlsx** » où <institut> désigne l'institut concerné, dans le répertoire commun « **BiblioMeter_Files/Impact Factor** » sous **S** ;
2. Ajoutez une feuille vierge après la feuille nommée < $n-2$ > (ex : en 2024, la dernière feuille sera nommée 2022) ;
3. Nommez la nouvelle feuille < $n-1$ > ;
4. Copiez la ligne des noms de colonnes de la feuille < $n-2$ > sur la 1^e ligne de la feuille < $n-1$ > ;
5. Changez le nom de la colonne nommée « IF $n-2$ » en « IF $n-1$ » (ex : en 2024, remplacez « IF 2022 » par « IF 2023 ») ;

La mise en forme de la feuille sera faite automatiquement par l'application lors de la mise à jour des facteurs d'impact. Si vous souhaitez mettre en place dès à présent la même mise en forme que les feuilles déjà existantes, effectuez les opérations suivantes :

1. Sélectionner la feuille < $n-2$ > ;
2. Sélectionnez l'ensemble du contenu de la feuille < $n-2$ > en cliquant dans l'angle en haut à gauche de cette feuille ;
3. Copiez le format de la feuille en cliquant sur le pinceau du menu « Accueil »
4. Sélectionner la feuille < $n-1$ > ;
5. Sélectionner l'ensemble du contenu de la feuille < $n-1$ > en cliquant dans l'angle en haut à gauche de cette feuille.

XVII. Choix de la période d'analyse

L'application est configurée pour traiter 6 corpus d'années consécutives à partir de l'année courante. Par exemple en 2024, les corpus qui seront traités par l'application sont ceux des années suivantes : 2019, 2019, 2020, 2021, 2022, 2024.

Pour pouvoir traiter **<n> années antérieures** à ces 6 années, il convient d'effectuer les opérations suivantes, dans le dossier de travail « **BiblioMeter_Files** » :

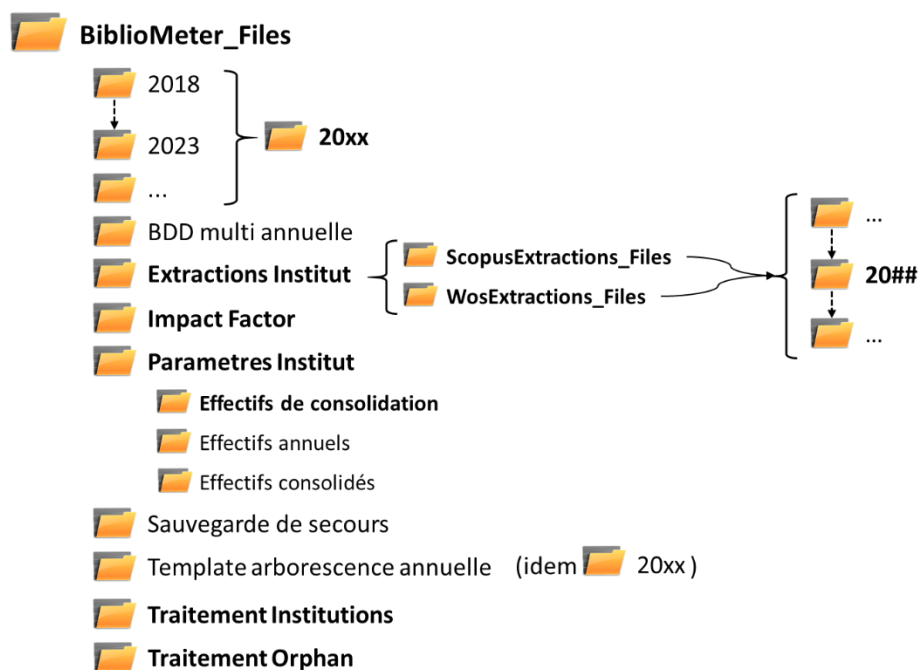
1. Créez les dossiers des **<n> années antérieures** s'ils ne sont pas disponibles ;
2. Mettez à jour les fichiers issus de l'extraction de WoS et Scopus dans le répertoire ad-hoc de ces dossiers ;
3. Déplacez les dossiers des **<n> années les plus récentes** dans un dossier de réserve ;
4. Lancez l'application ;
5. Effectuez le traitement ;
6. Remettez les dossiers des **<n> années les plus récentes** dans le dossier de travail.

Exemples :

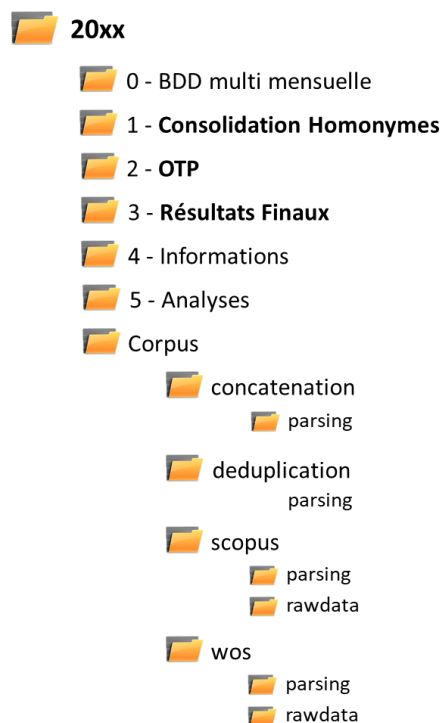
- pour traiter 2018, il suffit de mettre en réserve 2024 ;
- pour traiter 2016 et 2017, il faut mettre en réserve 2023 et 2024.

XVIII. Arborescence du dossier « BiblioMeter_Files »

1- Arborescence générale



2- Arborescence des dossiers annuels 20xx



En gras les dossiers où l'utilisateur a à intervenir manuellement.

Retour : IV

