

Tutoriel Transkribus CRHXIX

Tuto 4 : exporter les transcriptions

Par Lucie Slavik, stagiaire, en stage filé trois jours par semaine, 4 mai 2020 – 31 juillet 2020
Juin 2020

Matériel :

Export des transcriptions qui ont été faites dans le dossier BNF9enJPG (LP à Monseigneur Félix Dupanloup).

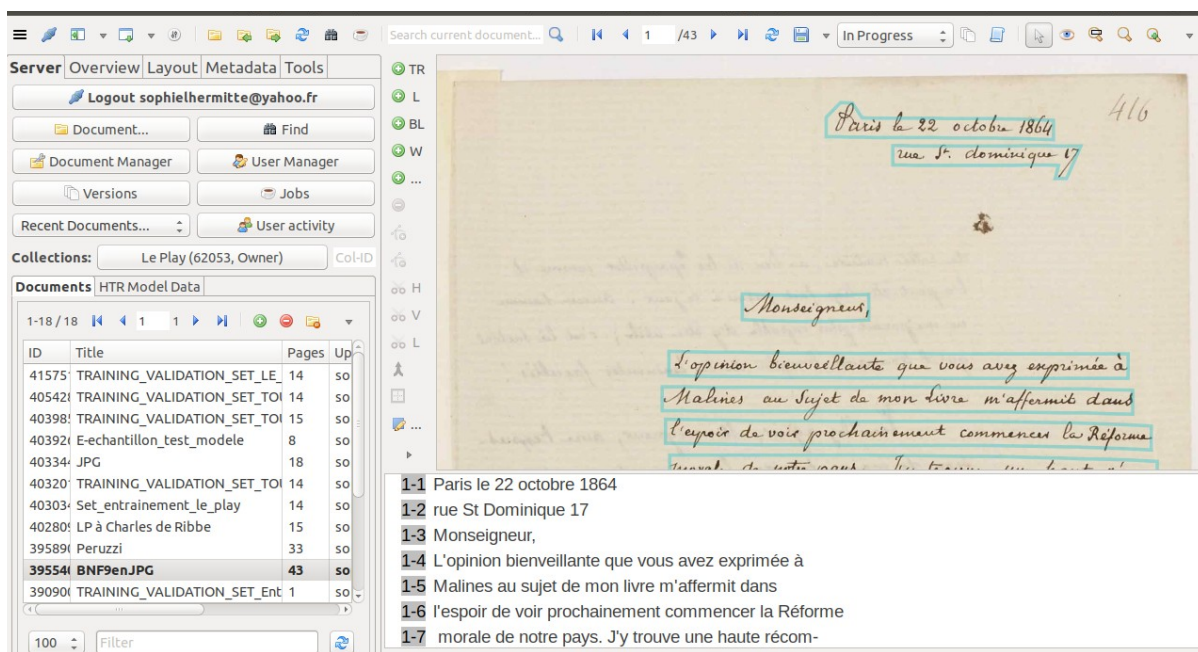
Tuto de base qui a servi pour la présente adaptation au CRHXIX : Comment_utiliser_Transkribus_-_en_10_étapes_ou_moins_with_Screenshots.pdf, disponible sur internet : <http://regis-schlagdenhauffen.eu/wp-content/uploads/2018/01/Comment-utiliser-Transkribus-%E2%80%93-en-10-%C3%A9tapes-ou-moins.pdf>

On peut exporter en plusieurs format.

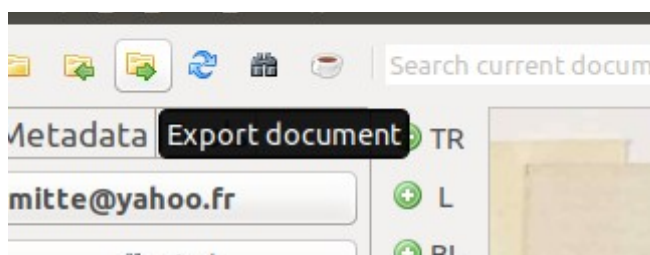
Aller sur Transkribus.

Choisir la collection de laquelle on veut exporter les textes

Ici donc, c'est la collection Le Play et plus spécifiquement le dossier BNF9enJPG.

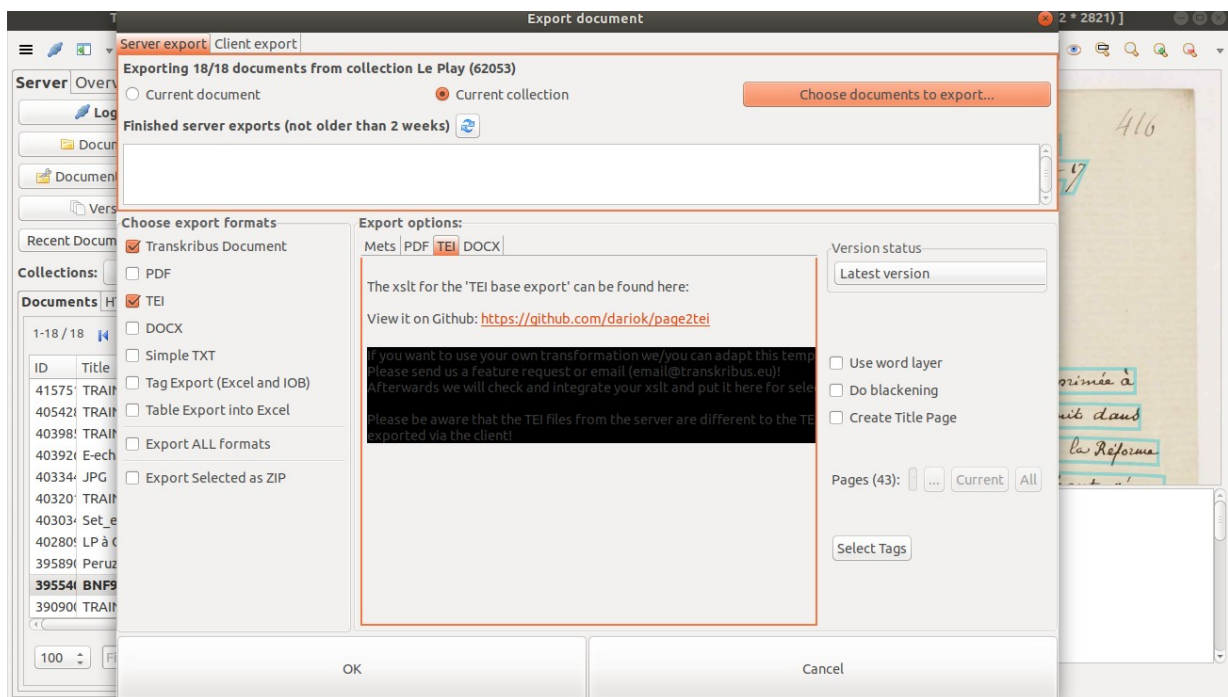


Pour commencer à exporter, cliquer en haut à gauche sur l'icône représentant un dossier avec une flèche verte qui va vers la gauche.

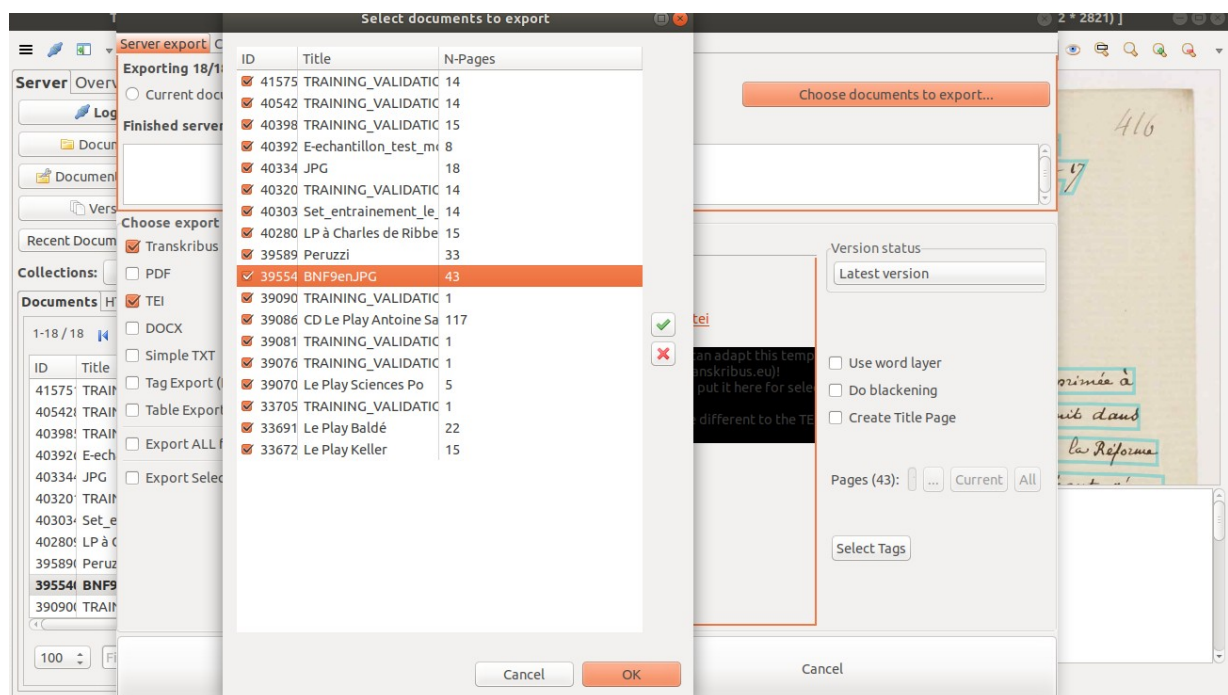


Cela ouvre une fenêtre avec deux onglets.

Sur le premier, Server export, choisir le format avec lequel on veut exporter les transcriptions. Ici, je clique sur TEI (pour les autres options, notamment celles indiquées à droite, comme « use word layer » etc., il faudrait que je me renseigne encore pour voir comment exporter le mieux) Préciser qu'on veut exporter la « current collection » pour avoir tout le dossier et non une seule page.



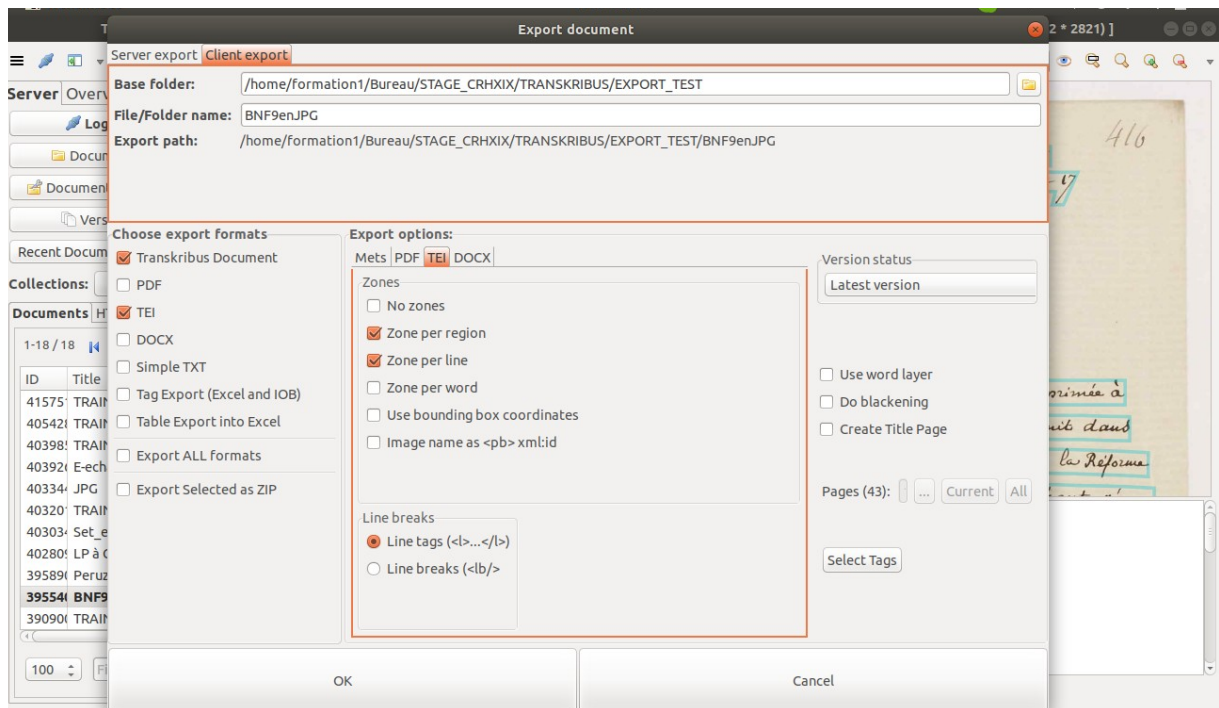
Cliquer sur « choose documents to export »



Une fenêtre s'ouvre. Sélectionner le dossier contenant les transcriptions qui nous intéressent et cliquer sur OK.

Puis aller sur l'autre onglet, Client export

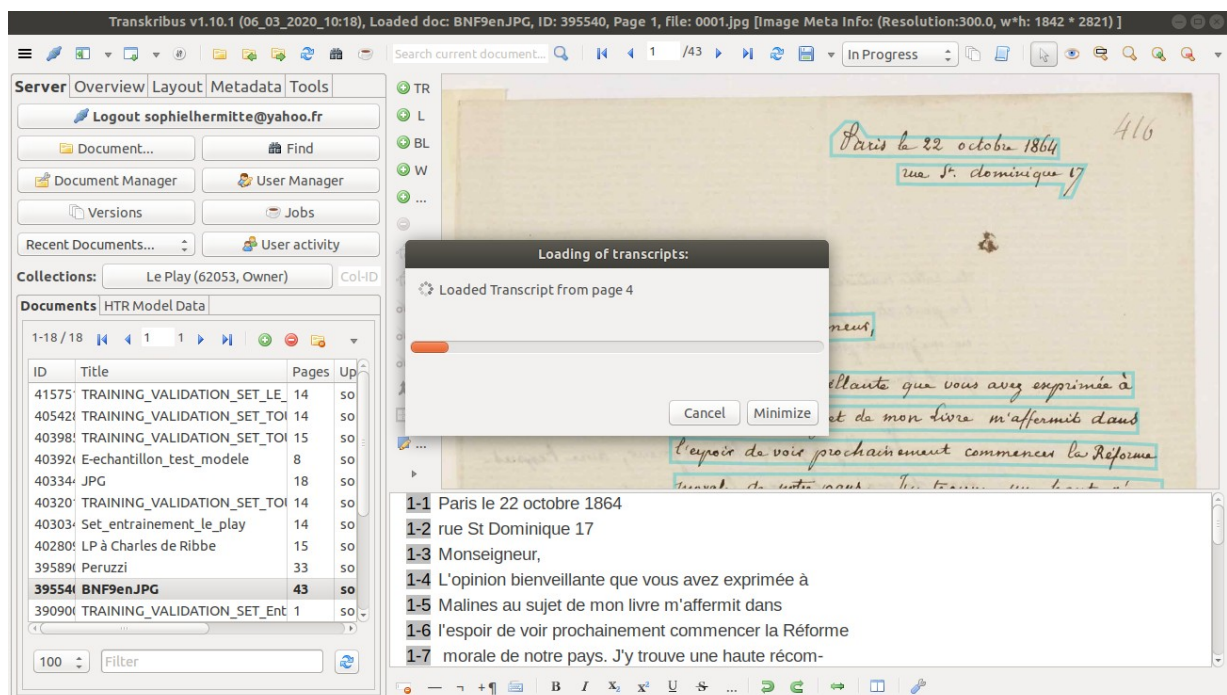
Et cliquer sur Choose documents to export...
Choisir le dossier qui nous intéresse et cliquer sur OK.

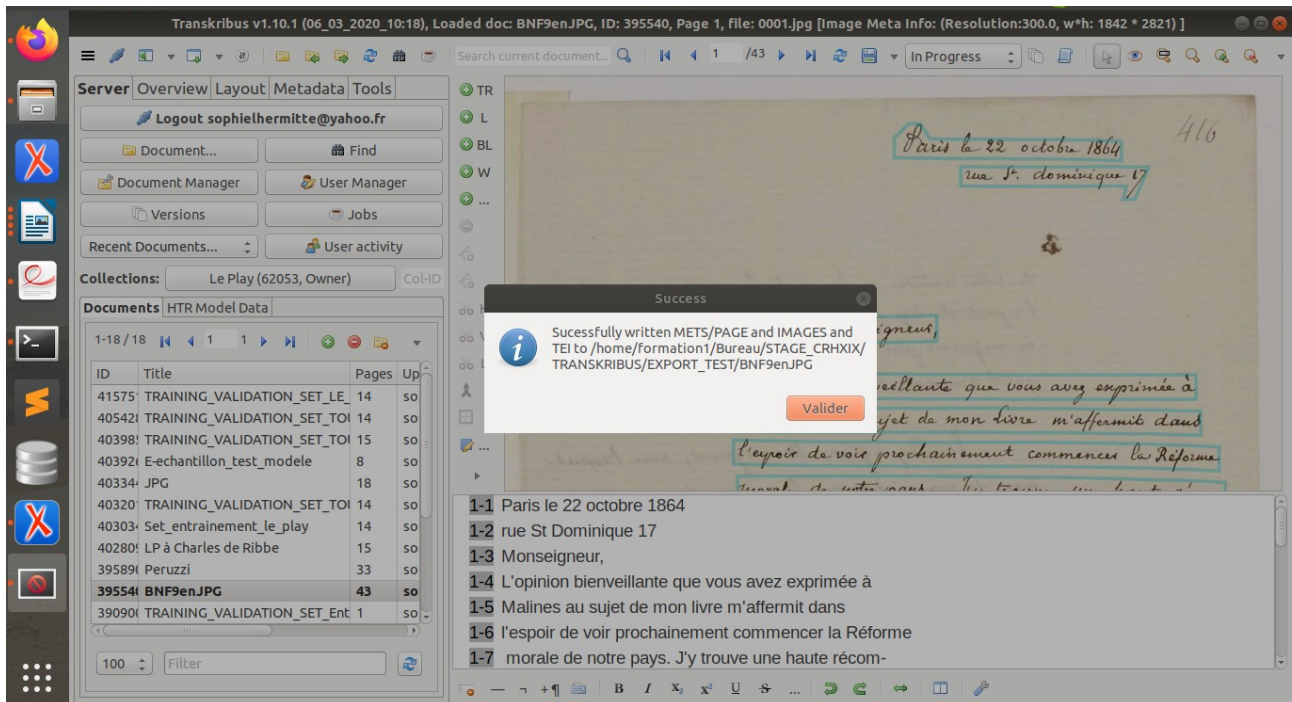


Choisir en cliquant sur l'icône représentant un dossier où l'on veut mettre l'export des transcriptions.

Pour les options d'export, il faudrait encore que je me penche plus sur la question.
Cliquer sur OK.

Et c'est parti ! L'exportation des transcriptions en XML TEI est en cours !





On peut alors aller dans le dossier où l'on a choisi d'exporter les transcriptions.
Il s'y trouve alors à la fois les manuscrits et leurs transcriptions, ce qui est extrêmement pratique !

Un inconvénient ?

Les fichiers XML sont générés par page et non par lettre... Voir comment faire pour le futur site... Normalement, il y a un fichier TEI par lettre, mais comment faire quand la lettre est sur plusieurs page ? Tout est expliqué dans l'ODD.

Voir si avec python, on pourrait transformer le XML de transkribus en un XML plus adapté à la correspondance, selon le schéma/modèle que j'aurai fait.
Ou le faire avec XSLT...

Questions à poser au tuteur de l'ENC.

IMPORTANT : Il vaut mieux mettre directement les archives de qualité que l'on utilisera effectivement pour le site sur Transkribus plutôt que de mettre des photos provisoires de mauvaise qualité et qui de toutes façons ne seront pas utilisées pour le futur site. Cela nous évitera le tracas de rechercher ensuite celles que l'on veut mettre effectivement.