



Entrée Est site François-Mitterrand -- Gilles Coulon / BnF

ÉDITER MICHEL FOUCAULT (1992–2021)

COLLOQUES

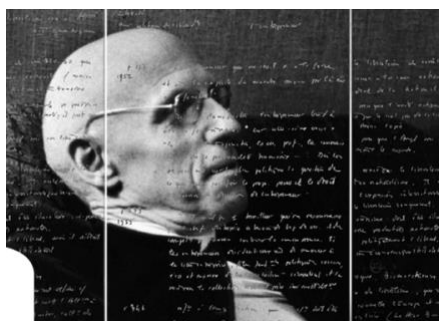
26 sep. 2020

9 h 30 - 18 h 30

François-Mitterrand
Grand auditorium

EXPERIMENTER TRANSKRIBUS

SUR LES FICHES DE LECTURE DE MICHEL FOUCAULT



Transkribus est une plateforme complète pour la reconnaissance, la transcription et la recherche automatisées de documents historiques manuscrits ou imprimés.

Elle facilite la lecture, l'enrichissement et l'exploration des textes. Le logiciel est développé par le [projet européen READ](#) qui souhaiterait révolutionner l'accès aux documents d'archives en développant des outils numériques innovants ([Transkribus](#), [ScanTent](#), [read@search](#)).

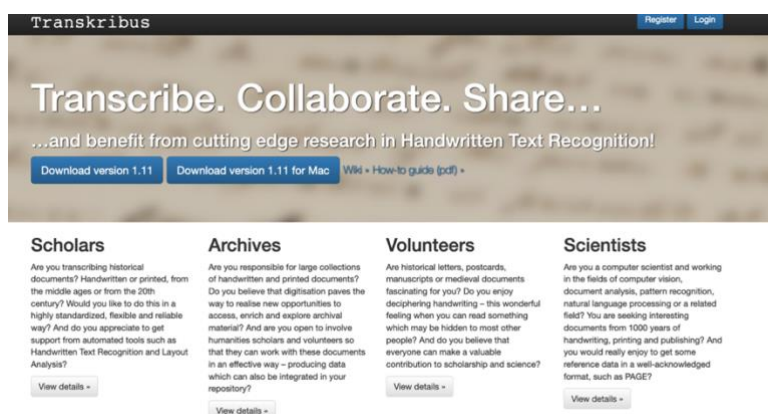
Dans le cadre du projet ANR [Foucault Fiches de lecture](#) (2017-2020), **l'équipe technique a entraîné un modèle à l'écriture manuscrite de Michel Foucault** (en segmentant et transcrivant manuellement 600 images de l'écriture du philosophe) puis a appliqué ce modèle HTR (reconnaissance automatique de caractère) pour transcrire automatiquement le reste du corpus (plus de 20 000 fiches manuscrites collectées pendant plus de 30 ans pour la préparation des cours, des interventions ou des ouvrages du philosophe). Les résultats obtenus par le projet sont très encourageants, puisque 95% des caractères sont reconnus automatiquement sans erreur par le logiciel Transkribus. Nous pouvons encore améliorer le modèle en corrigeant les transcriptions automatiques, afin d'entraîner un nouveau modèle encore plus performant. Il s'agit d'un **processus itératif**.

Ces **transcriptions automatiques** facilitent la lecture difficile des fiches de lecture de M. Foucault, font gagner du temps sur la transcription (40 min / fiche pour une transcription manuelle ; quelques minutes pour un lot de fiches !) et permet une analyse automatisée du corpus (indexation, recherches exactes ou approximatives, fouille de texte, etc.).

Néanmoins, les transcriptions doivent être corrigées et validées par des spécialistes de Foucault avant d'être mises en ligne.

S'enregistrer sur la plateforme et télécharger Transkribus :

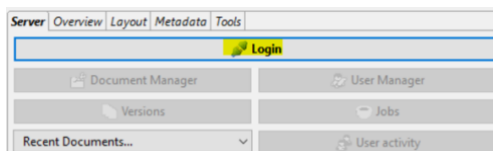
- Vous pouvez **consulter les guides** écrits et vidéos réalisés par l'équipe de Transkribus sur leur wiki : https://transkribus.eu/wiki/images/7/77/How_to_use_TRANSKRIBUS_-_10_steps.pdf (en anglais) ; https://transkribus.eu/wiki/images/1/1d/Comment_utiliser_Transkribus_%E2%80%93_en_10_%C3%A9tapes_ou_moins_with_Screenshots.pdf (en français).
- Vous pouvez aussi consulter le **tutoriel en français établi par Emmanuelle Perrin** (Archéorient, Université de Lyon), réalisé en novembre 2019 : https://hal.archives-ouvertes.fr/hal-02472234/file/Tutoriel_Transkribus_BBM.pdf
- S'**enregistrer** sur la plateforme et **télécharger** le logiciel Transkribus : <https://transkribus.eu/Transkribus/> ; **Procédure pas à pas** aux adresses suivantes : <https://transkribus.eu/wiki/index.php/Registration> et https://transkribus.eu/wiki/index.php/Download_and_Installation



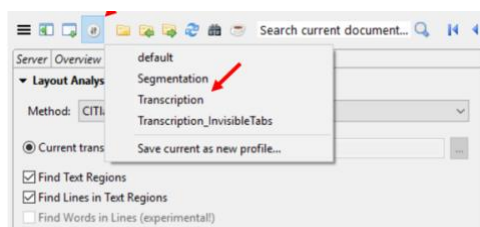
- **Envoyer ensuite un message** à marie-laure.massot@ens.fr et vincent.ventresque@ens-lyon.fr en communiquant votre nom de connexion et votre adresse mail de connexion sur la plateforme Transkribus.
- L'équipe technique du projet FFL vous donnera un **accès à un corpus test**.

Scénario pour tester le logiciel sur le corpus des fiches de lecture :

- Connectez-vous au serveur Transkribus avec vos identifiants et ouvrez la collection « tests BNF » : Login / Serveur / double-cliquer sur la collection « tests BNF » /

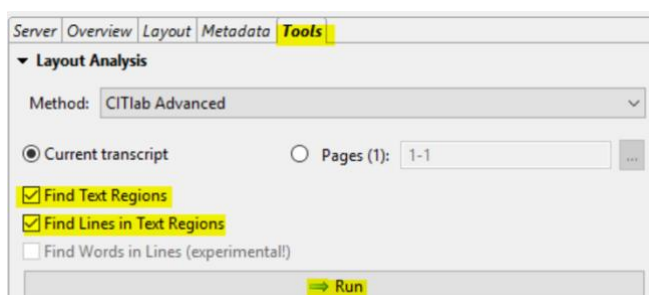


- Dans la collection « tests BNF », consultez la première fiche et la transcription correspondante pour vous familiariser avec l'écriture de Michel Foucault : double-cliquez sur B1_T0000003.jpeg / Cliquez sur le bouton « Visualiser les profils (icône flèches verticales vers le bas et vers le haut) dans le menu principal et sélectionnez Vue « Transcription ».



- Segmenter la fiche de lecture n°10 de la boîte 38 (B38_T0000020 et B38_T0000021) :

Pour transcrire automatiquement une fiche de lecture avec le modèle HTR Foucault¹, les fiches doivent être segmentées (faire un lien direct entre les images du document et le texte transcrit correspondant). Ce processus de division d'une page en régions, lignes et lignes de base est appelé **segmentation** et se fait de manière semi-automatique sur la plate-forme : Cliquez sur le bouton « Visualiser les profils » dans le menu principal et sélectionnez Vue « Segmentation » (cf. Capture d'écran ci-dessus). Pour détecter automatiquement les régions de texte, allez dans l'onglet « Outils » et cliquez sur « Détecter les lignes et les lignes de bases ». Vous pouvez traiter plusieurs pages en cochant « Pages » et en indiquant les pages que vous souhaitez segmenter automatiquement.

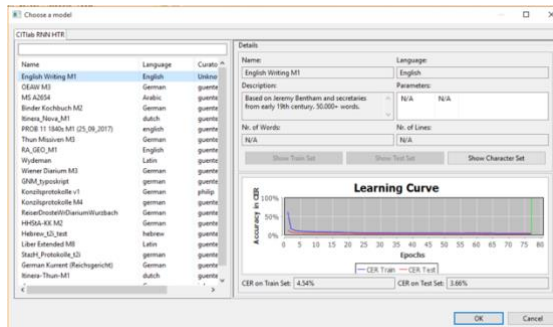
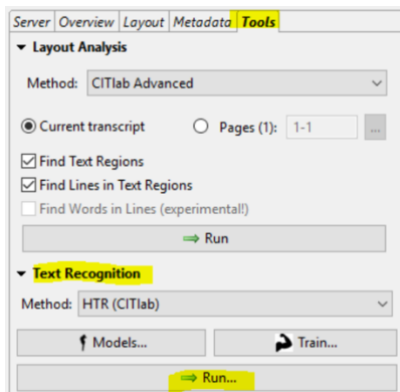
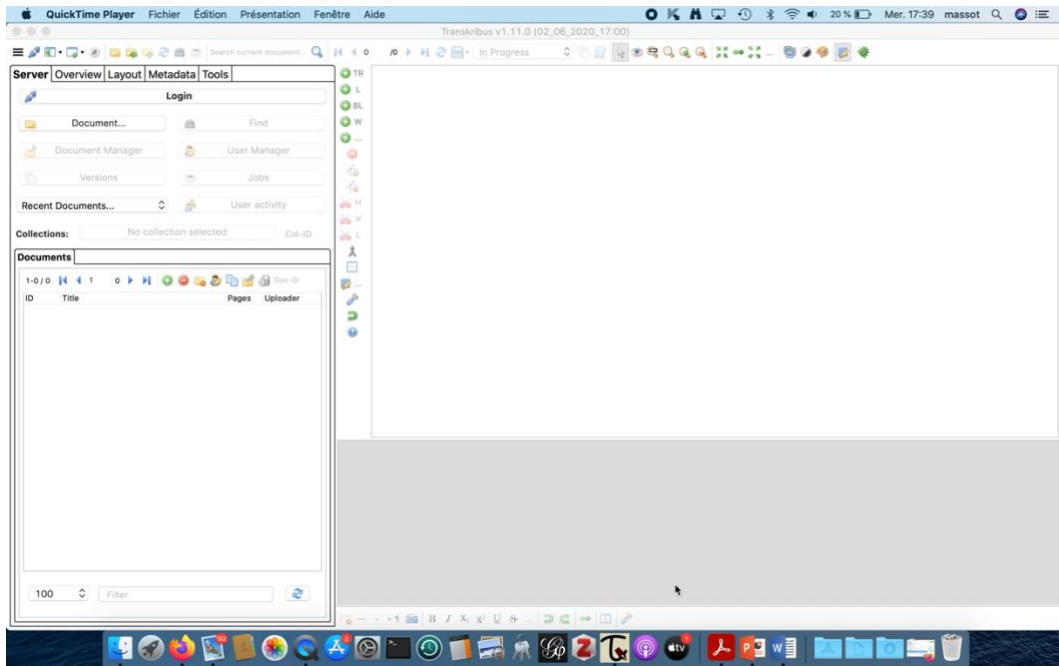


- Transcrire automatiquement la fiche segmentée (n°10 de la boîte 38) avec le dernier modèle HTR pour l'écriture manuscrite de Michel Foucault et consulter le taux d'erreur sur caractère du modèle :

Le **modèle HTR Michel FoucaultM3 +** qui a été formé par l'équipe technique du projet FFL peut être appliqué via le module « Exécuter la reconnaissance de texte » dans l'onglet « Outils ». Vous pouvez sélectionner une ou plusieurs pages de vos documents et démarrer le traitement HTR. (Vue "Transcription" / Tools / Models /

¹ Modèle de reconnaissance automatique de l'écriture manuscrite de Michel Foucault.

Choose MichelFoucaultM3+ / Run). Consultez la vidéo "Transkribus_automatique.mov" pour vous guider :



- Faire une **recherche de mots clés** (Keyword Spotting, **KWS**²) en recherchant le mot « ange » dans la fiche (n°10 de la boîte 38) segmentée et transcrite automatiquement. (Search (icône jumelles) / Keyword "ange" /Search /). Remarquons que la recherche par KWS identifie bien 3 images correspondantes au mot clé "ange", alors que la transcription automatique est fautive pour 2 des 3 traductions : "anges", Arge" et "Argens". Vous pouvez consulter la **vidéo "Transkribus_KWS.mov"** pour vous guider :

The screenshot shows the Transkribus v1.11.0 interface. The main window displays a document with handwritten text in French, which has been segmented and transcribed. The text is: "L'union c'icution avec Satan. 10", "La c'ication de p'h avec les anges est "dérivée de la c'auté de p'ître que m remarquons en ces 2 natures:", "Au contraire "l'association de p'h chassé du Paradis avec l'Arge Garre du liel est fondée ou la condition des leur c/n accidants".

Below the document, a list of search results is shown for the keyword "ange":

Search ID	Text
1-1	10
1-2	E'heest en c' icution pvec Satan.
1-3	La c'ication des fh avec les anges
1-4	est "dérivée de la c'auté de p'ître que
1-5	pas remarquons en ces 2 natures.
1-6	Au contraire "l'association de fh chassé
1-7	du Paradis avec l'Arge Garre du liel
1-8	est fondée ou la condition des leur c/n accidants
1-9	21)
1-10	AGandonne de Dieu, et séparé des Argens

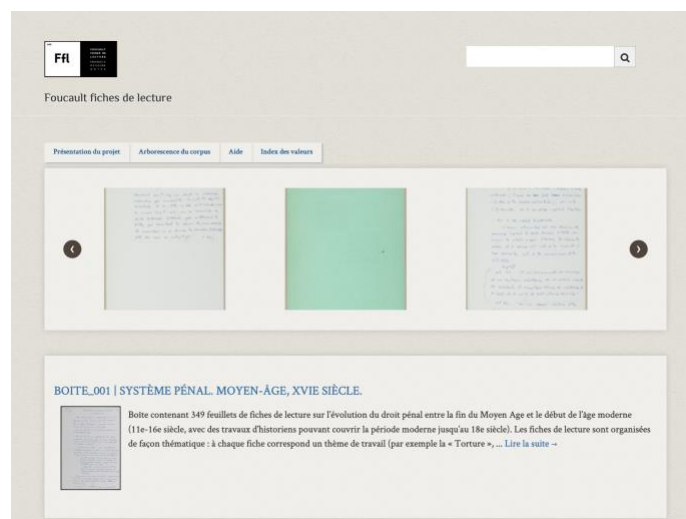
The interface also shows a sidebar with document management options and a table of documents in the background.

² Outre la recherche plein texte (recherche exacte d'un terme) et la recherche floue (recherche approximative sur le texte), la plateforme Transkribus propose un outil de recherche innovant pour traiter des manuscrits historiques, le Keyword Spotting (KWS). La recherche approximative KWS développée par les informaticiens de l'équipe de Transkribus permet au logiciel de comparer des images de mots et ainsi de retrouver, même dans le cas d'une transcription fautive, les différentes occurrences de ce mot dans les images.

En savoir plus sur le projet Ffl :

Carnet de recherche du projet : <https://ffl.hypotheses.org/>

Consulter les fiches de lecture de Michel Foucault sur la plateforme publique FFL-Eman :



<http://eman-archives.org/Foucault-fiches/>

Compte-rendu, article et communications dans le cadre du projet Ffl

[Transcrire les fiches de lecture de Michel Foucault avec le logiciel Transkribus : compte rendu des tests.](#)

Juin 2018 : Compte-rendu des tests consultable dans HAL.

[Transcribing Foucault's handwriting with Transkribus – Report](#)

Nov. 2018 : Communication au deuxième congrès des utilisateurs de Transkribus (Vienna), Autriche.

[Transcribing Foucault's handwriting with Transkribus. Jdmdh:5043 – Journal of Data Mining & Digital Humanities, March 5, 2019, Atelier Digit Hum.](#)

Mars 2019 : Article en anglais publié dans jdmdh.


[Transcrire automatiquement les fiches de lecture de Michel Foucault avec Transkribus : retour d'expérience.](#)

Avr. 2019 : Communication à l'atelier international de travail EMAN-latex (Las Palmas Canaries), Espagne.

« Lire les fiches de lecture avec Transkribus »

Sept. 2020 : Communication au colloque de clôture du projet ANR FFL – 1ère Partie Fiches et fichiers. L'archive Foucault à l'ère du numérique (IMEC), Abbaye d'Ardenne.

En savoir plus sur le logiciel Transkribus :



Ressources

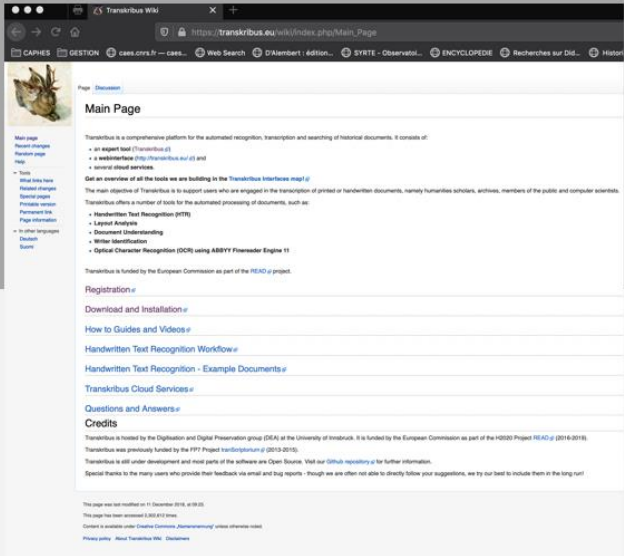
Contacter l'équipe :
TRANSKRIBUS Team at University of Innsbruck
email@transkribus.eu

S'enregistrer sur la plateforme et télécharger le logiciel Transkribus
<https://transkribus.eu/Transkribus/>

Guides d'utilisation et tutoriels sur Transkribus Wiki
https://transkribus.eu/wiki/index.php/Main_Page

https://transkribus.eu/wiki/images/7/77/How_to_use_TRANSKRIBUS_-_10_steps.pdf
https://transkribus.eu/wiki/images/1/1d/Comment_utiliser_Transkribus_%E2%80%93%93_en_10_%C3%A9tapes_ou_moins_with_Screenshots.pdf

Vous pouvez aussi consulter le [tutoriel en français établie par Emmanuelle Perrin](#) (Archéorient, Université de Lyon), réalisé en novembre 2019 : https://hal.archives-ouvertes.fr/hal-02472234/file/Tutoriel_Transkribus_BBM.pdf



Retrouvez les vidéos "Transkribus_automatique.mov" et "Transkribus_KWS.mov" sur le **Carnet FFL**
Scénario réalisé dans le cadre du projet ANR *Foucault Fiches de lecture* (ANR-17-CE38-001)
Marie-Laure Massot, Jean-Philippe Moreux et Vincent Ventresque pour tester le logiciel
Transkribus sur les fiches de lecture de Michel Foucault. Version du 19/09/2020.



Interdiction de diffuser des transcriptions sans l'accord préalable des ayants droits de Michel Foucault.

