



Mettre en ligne, annoter et exploiter les fiches de lecture de Michel Foucault

Marie-Laure Massot

► To cite this version:

Marie-Laure Massot. Mettre en ligne, annoter et exploiter les fiches de lecture de Michel Foucault. Master. Atelier autour des archives, Centre documentaire du CAPHES, France. 2023. hal-04057849

HAL Id: hal-04057849

<https://hal.science/hal-04057849>

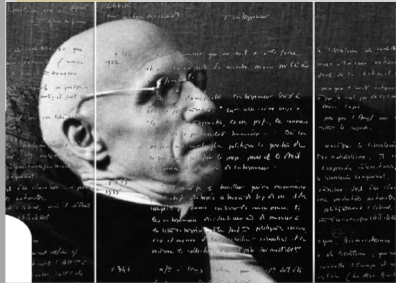
Submitted on 4 Apr 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0
International License



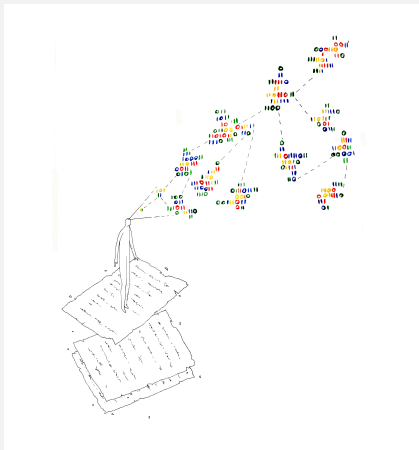
Mettre en ligne, annoter et exploiter les fiches de lecture de Michel Foucault

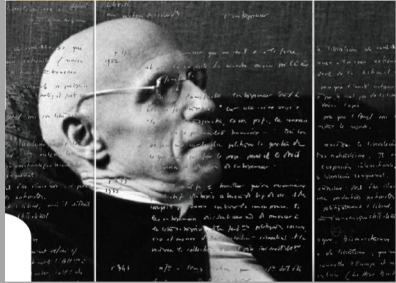
Marie-Laure Massot (ingénieure d'étude SHS)

CAPHÉS, Unité d'appui et de recherche – CNRS/ENS PSL

Atelier autour des archives

3/04/2023 ENS PSL





Mettre en ligne, annoter et exploiter les fiches de lecture de Michel Foucault

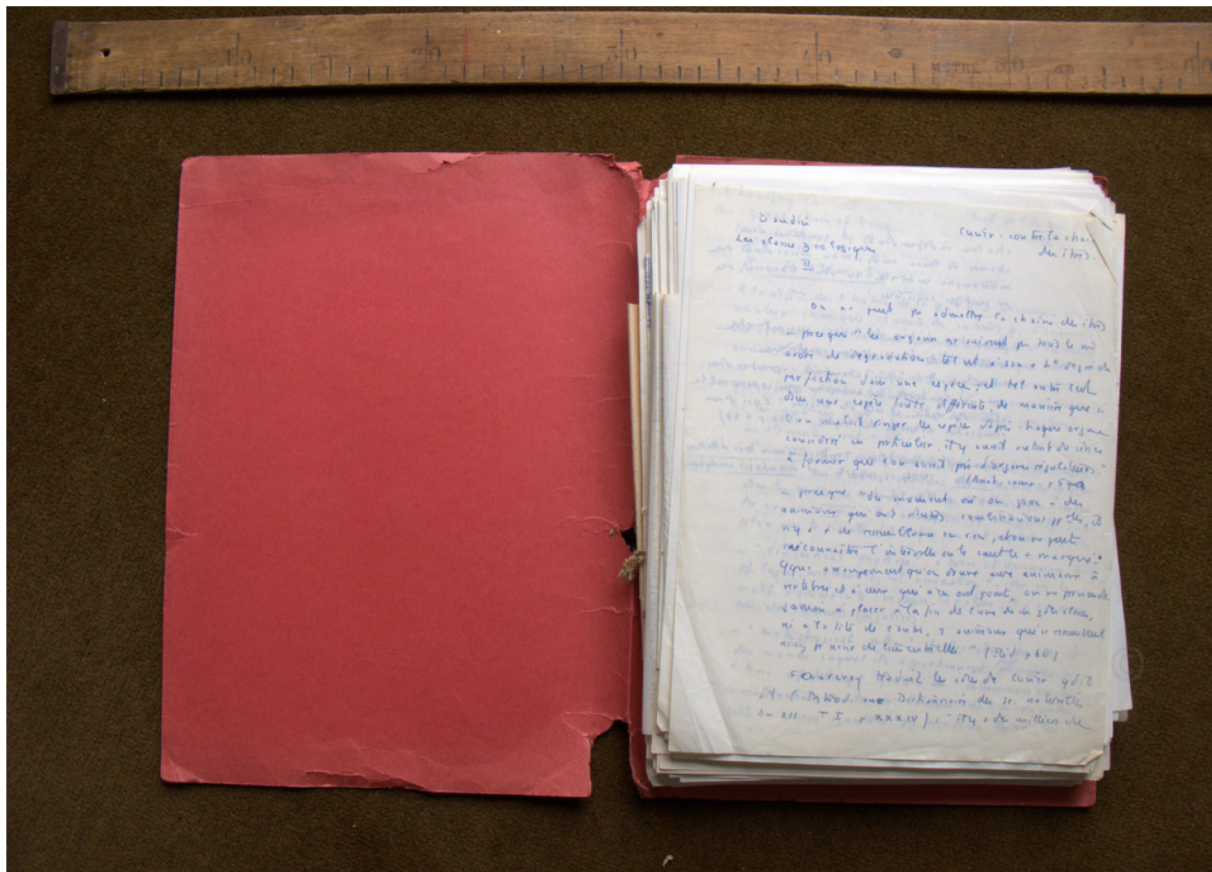
Résumé :

Avec près de **20 000 feuillets numérisés et mis en ligne**, la plateforme Eman a permis au projet Foucault fiches de lecture de diffuser très largement le fonds Foucault de la BnF. Nous verrons **comment cet outil modulaire a pu être adapté**, moins pour la production des données de description et d'indexation que **pour leur exposition et exploitation**. En particulier, nous présenterons la **chaîne de traitement et l'articulation avec les autres outils du projet FFL**, le prototype de plateforme collaborative et Transkribus, et les **développements informatiques réalisés pour Eman au cours du projet**.



Mettre en ligne, annoter et exploiter les fiches de lecture de Michel Foucault

Les archives de Michel Foucault (département des manuscrits de la BnF)



triangle
UMR 5206

{ BnF



ANR-17-CE38-0001

Mettre en ligne, annoter et exploiter les fiches de lecture de Michel Foucault

Ma gestion du projet recherche collaboratif FFL – ANR 2017-21 / Révolution numérique



- **Coordinatrice pour le partenaire ENS PSL**
- **Contribution à la réflexion sur les outils et méthodes** à mettre en œuvre
- **Expérimenter et transférer mes compétences** du système expert *Transkribus*
- **Mise en ligne** de la plateforme FFL-Eman
- **Valorisation des réalisations et des résultats obtenus** par l'équipe pluridisciplinaire (communications, séminaires, articles, formations)

Mettre en ligne, annoter et exploiter les fiches de lecture de Michel Foucault

Le prototype FFL : description, annotation, exploration

1/2

1

2

Rechercher contenu

Niveau associé

1

2

Rechercher d'autres fiches de la boîte :

genre

←

→

Duby
Eco rurale
17

Exploitation des paysans au XI^e-XII^es. 2
La justice / fisc.

- Au cours du XI^es, exploitation des paysans, diabolisation etc. Dieu enrichit les seigneurs. Avec la même technique, les seigneurs ont moins besoin d'h. de leur cour, et donc moins. Dieu au pouvoir en h., il préfère 1 seigneur sur 2 seigneurs etc.

Bloc-notes / mémo

File - Edit - View - Format -

1

2

Aucun contenu enregistré dans le bloc-note.

Enregistrer la bloc-notes

Powered by Trivium

Bloc notes personnel de Marie Laure Massot [?]

Décrire / annoter la fiche [b001_f0002]

Type de l'annotation : 1.1 Descripteur (test-04-april) Valider type choisi [?]

Description et annotations de la pièce n°b001_f0002 [Boîte_001]

DescriptionArchivistique

MentionDePersonne

Duby, Georges

[?] action

TermeExtrait

exploitation

[?] action

TermeExtrait

fisc

[?] action

TermeExtrait

justice

[?] action

TermeExtrait

rural

[?] action

TermeExtrait

économie

[?] action

TitrePiece

Duby, Eco rurale II. | Exploitation des paysans au XI^e-XII^es. La justice comme fisc.

[?] action

AnalyseScientifique

Descripteur

préparation de Surveiller et punir

[?] action

Descripteur

paysannerie

[?] action

NoteEditoriale

Avertissement

Reprendre les références de cette fiche

[?] action

Type de la pièce

Type de la pièce : fct.FicheLecture. Sélectionner un nouveau type dans la liste déroulante puis valider en cliquant sur le bouton.

Procédure

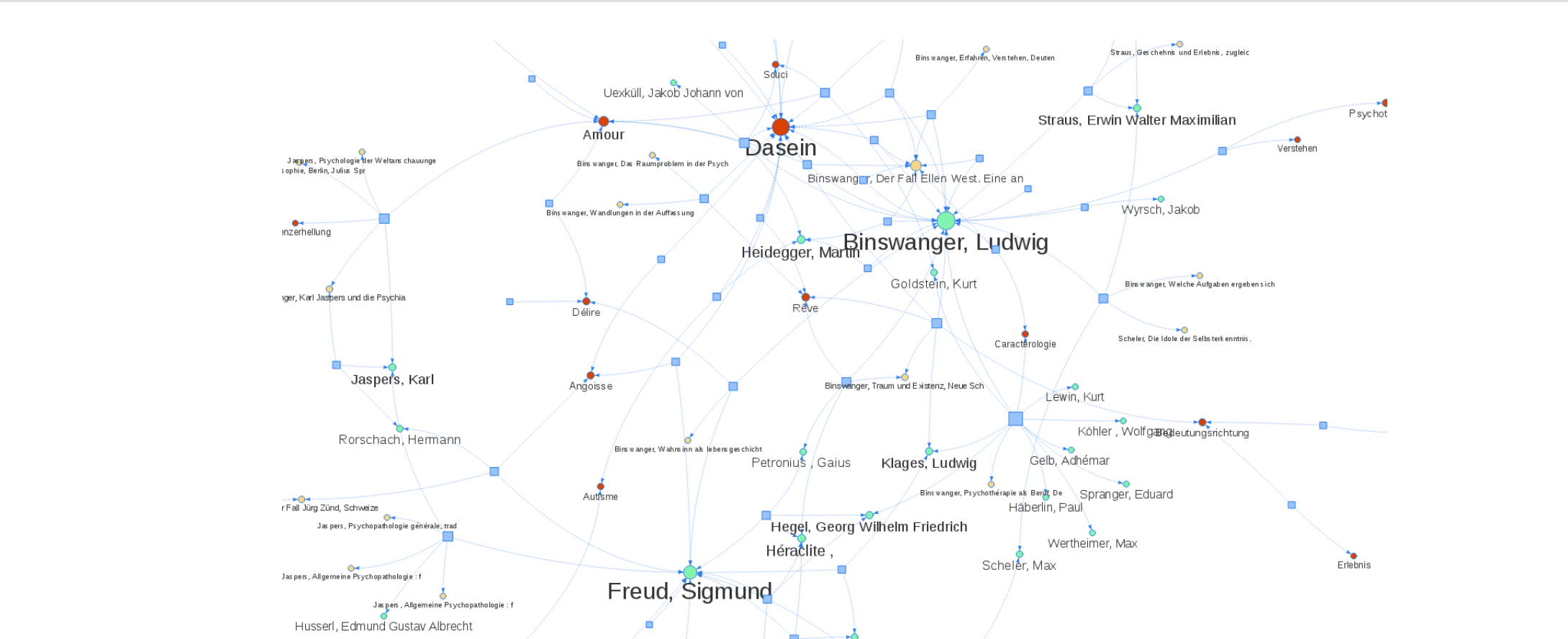
Changer le type de la pièce [?]

Contrôle des requêtes et données de la page

Afficher

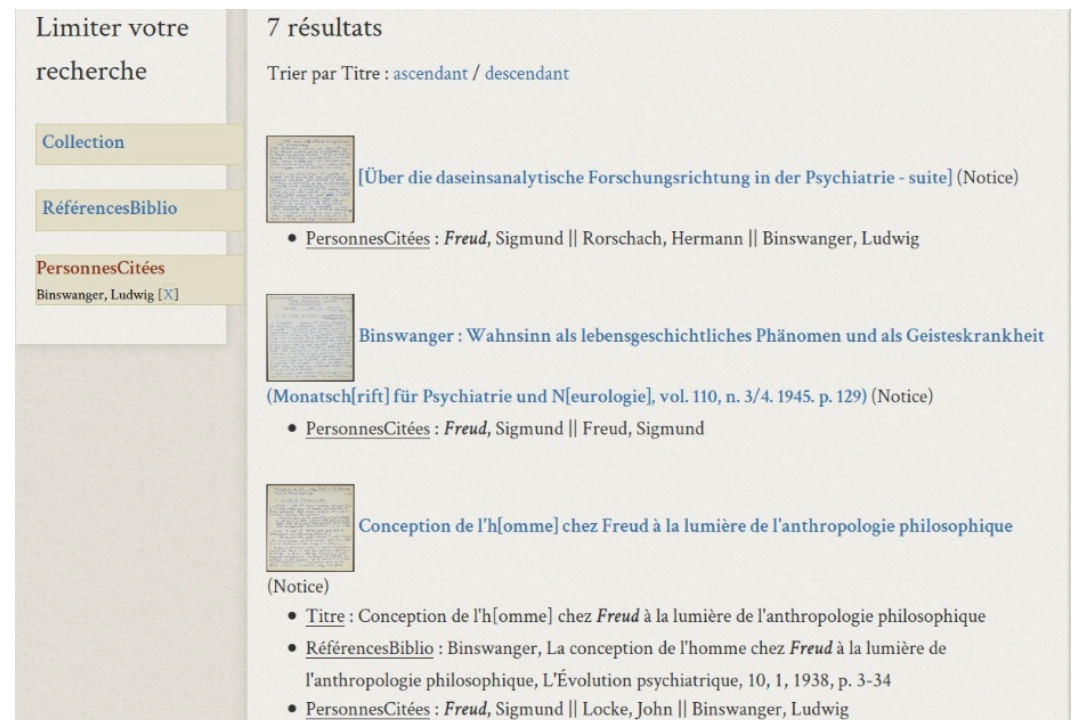
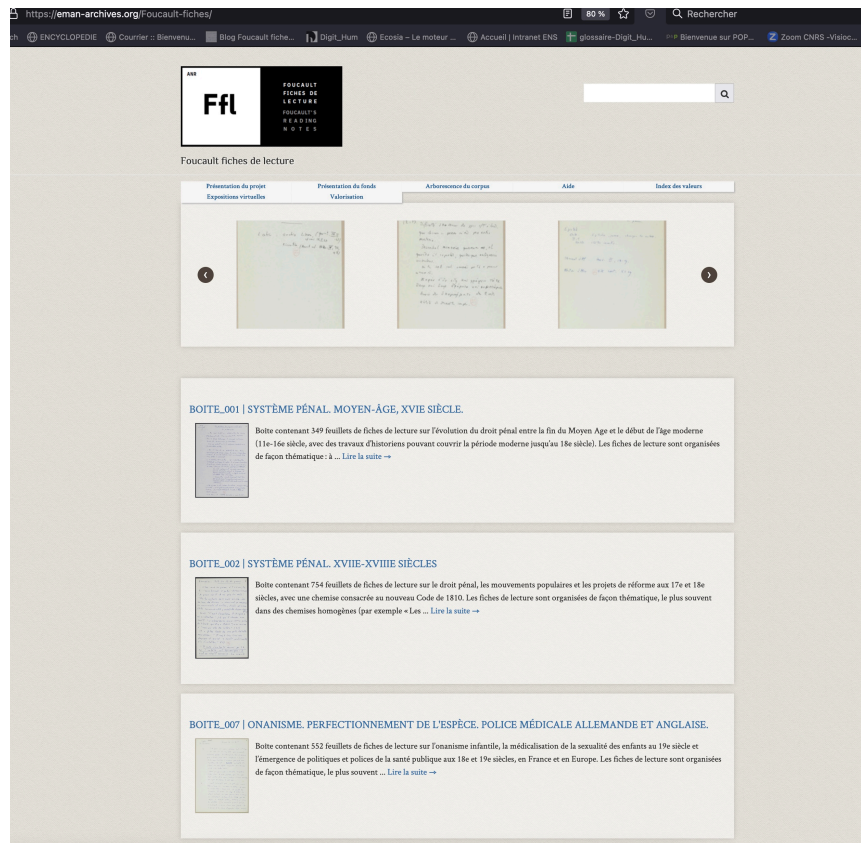
Mettre en ligne, annoter et exploiter les fiches de lecture de Michel Foucault

L'outil de cartographie : navigation dans les contenus produits par l'équipe



Mettre en ligne, annoter et exploiter les fiches de lecture de Michel Foucault

La plateforme publique FFL-EMAN basée sur Omeka

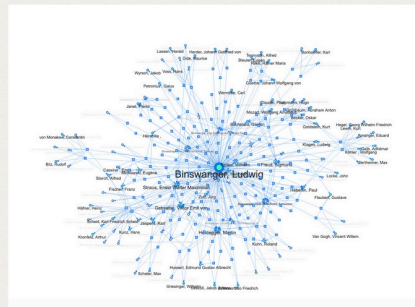


Mettre en ligne, annoter et exploiter les fiches de lecture de Michel Foucault

L'exposition virtuelle : une ballade dans le corpus

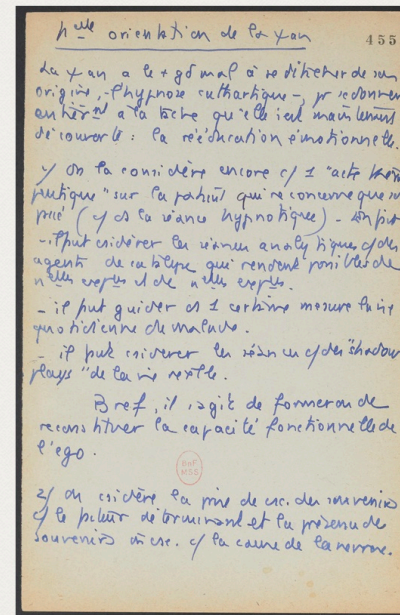
Des recherches ciblées

Si certaines parmi ces références bibliographiques restent de seconde main dans les écrits de Foucault, grâce aux fiches de lecture nous pouvons désormais découvrir les textes qui avaient attiré le plus son attention, les passages qu'il avait soulignés, les concepts sur lesquels il s'était arrêté. En général, grâce aux fiches de lecture j'ai pu accéder à un très grand nombre d'informations qui se sont révélées fondamentale pour l'édition du manuscrit *Binswanger et l'analyse existentielle*, notamment pour la rédaction des notes bibliographiques et des notes critiques. Ces informations concernent notamment les **données bibliographiques**, aussi bien que les **notions et concepts** présents dans le texte de Foucault, qui fonctionnent comme des sources lexicales et conceptuelles dans le travail d'édition.



[cliquer sur l'image pour accéder à la carte interactive]

Or, comment faire pour retrouver les informations dont l'éditeur a besoin parmi le millier de feuillets consultés ? C'est là qu'intervient la plateforme FFL, qui fonctionne comme une base de données constituée à partir des informations que j'avais insérées préalablement. Autrement dit, j'ai adapté l'outil à mes problématiques spécifiques.



Fiche de lecture "Nouvelle orientation de la psychanalyse" B038--f0455

Voici quelques exemples pour mieux comprendre comment j'ai travaillé :

- Dans le manuscrit, de nombreuses données bibliographiques sont lacunaires (incomplètes ou manquantes). Foucault renvoie souvent à une note en bas de page, mais sans fournir ensuite aucune référence. C'est le cas, par exemple, du psychanalyste américain d'origine hongroise Franz Alexander, que Foucault cite dans son manuscrit à propos des « *certaines tendances récentes de la méthode thérapeutique chez les analystes* ».

Pour chercher la référence que Foucault pouvait avoir en tête lorsqu'il a rédigé ce passage, j'ai lancé une recherche par auteur dans la plateforme et j'ai obtenu toutes les occurrences du nom « Alexander » présentes dans les fiches de lecture numérisées dans le cadre du projet FFL. Parmi les diverses informations obtenues, en particulier, il y a une fiche de lecture intitulée « Nouvelle orientation de la psychanalyse » :

- Fiche de lecture « Nouvelle orientation de la psychanalyse » b038--f0455

J'ai donc sélectionné la fiche (qui se trouve dans la boîte 38) et j'y ai retrouvé le même passage mentionné par Foucault dans son manuscrit, avec en plus l'indication bibliographique manquante : *Psychoanalytic Therapy. Principles and Application*, New York, The Ronald Press Company, 1946.

Mettre en ligne, annoter et exploiter les fiches de lecture de Michel Foucault

Focus sur l'expérimentation Transkribus - Présentation



Développé dans le cadre du **projet européen READ** (H2020), le système expert Transkribus promettait de révolutionner le travail sur les documents d'archives grâce à l'intelligence artificielle et à ses **fonctionnalités de reconnaissance automatique d'écriture manuscrite** (Handwritten Text Recognition, **HTR**).

Une collaboration avec l'équipe du projet READ, entamée en 2017, nous a permis **d'expérimenter Transkribus sur les fiches de lecture de Michel Foucault**, pour voir dans quelle mesure ce logiciel pouvait nous aider à « lire » et à indexer les manuscrits du philosophe.

Le projet READ (*Recognition and Enrichment of Archival Documents*) a été financé par la commission européenne dans le cadre du programme Horizon 2020 (2016-2019). A la suite de ce financement européen, l'équipe a créé **une société coopérative européenne** pour continuer de soutenir et de développer la plateforme Transkribus : <https://readcoop.eu/>

! Certains services sont devenus payants !

<https://readcoop.eu/transkribus/>

Mettre en ligne, annoter et exploiter les fiches de lecture de Michel Foucault

Focus sur l'expérimentation Transkribus - Pourquoi transcrire de simples fiches de lecture ?

Transkribus v1.3.7 (23_11_2017_14:37), Loaded doc: Fonds_Foucault_BnF_boite-1-jpeg, ID: 23825, Page 3, file: T0000003.jpeg [Image Meta Info: (Resolution:400.0, w*h: 3404 * 4766)] [current line: w*h: 2346 * 126]

Search current document...

Server Overview Layout Metadata Tools

Logout marie-laure.massot@ens.fr

Document Manager User Manager

Versions Jobs

Recent Documents... User activity

Collections:

test_foucault_LBF (4016, Owner)

ID	Title	Pages
43640	boite22	1140
42205	boite20	1225
35689	boite51_diff	628
23825	Fonds_Foucault_BnF_boite-1-jpeg	698
8619	icp1	2

1-5 / 5

100

Region: 3 / 6

Handwritten text (transcription):

July
Eco-rumb
17

Exploitation du paysan au XI^e XII^e s.
Le jour h... / h...

- Au cours du XI^e s, xtiplication des censires,
défrichement etc. D'ou'enrichisse^{nt} des seigneurs.
Avec le progrès technique, les seigneurs ont moins
besoin d'h. ds leur cour, c/ domestiques.
D'ou'au XI^e s, il y a une préférence pour le travail sur la
parcelle etc. etc.

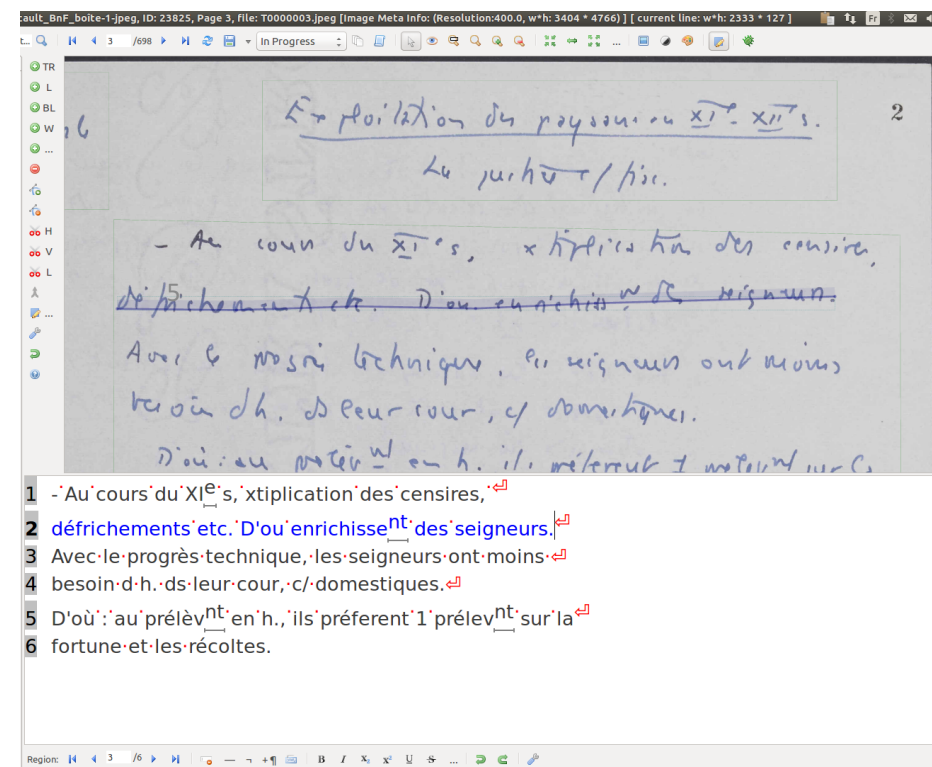
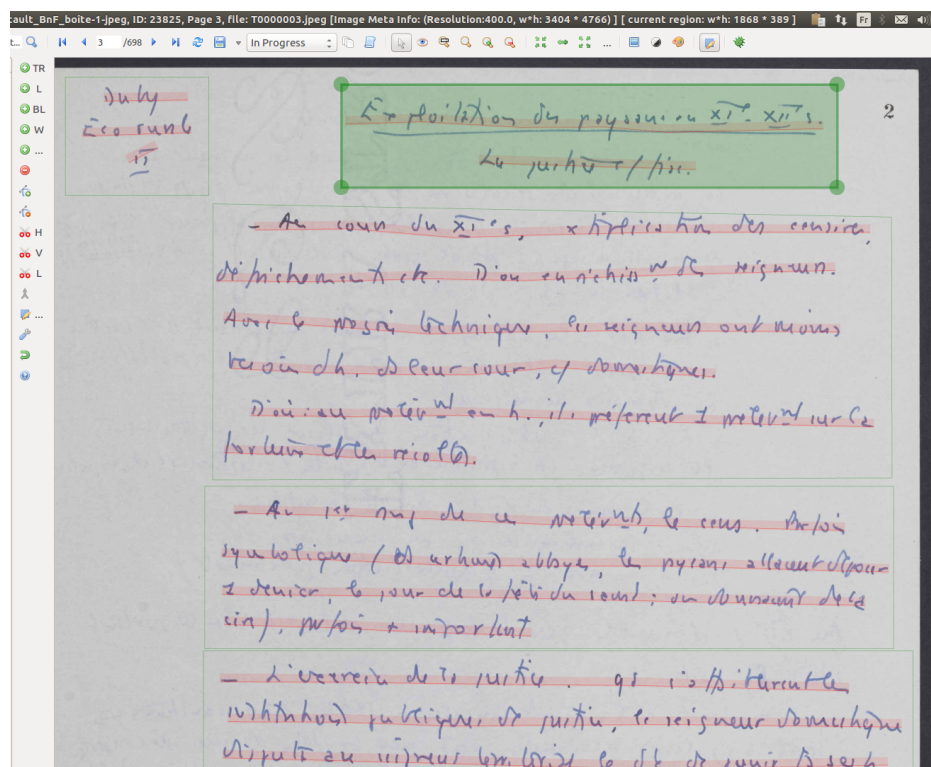
- Au 1^{er} s. du XI^e s, le cours. Alors
symbolique (de l'homme aboyé, le paysan, alléant d'après
et de l'homme, le paysan de la terre du seigneur; un d'un autre d'après
cin), parfois important



1. Pourquoi transcrire ?
2. Créer des données d'apprentissage
3. Résultats et perspectives

Mettre en ligne, annoter et exploiter les fiches de lecture de Michel Foucault

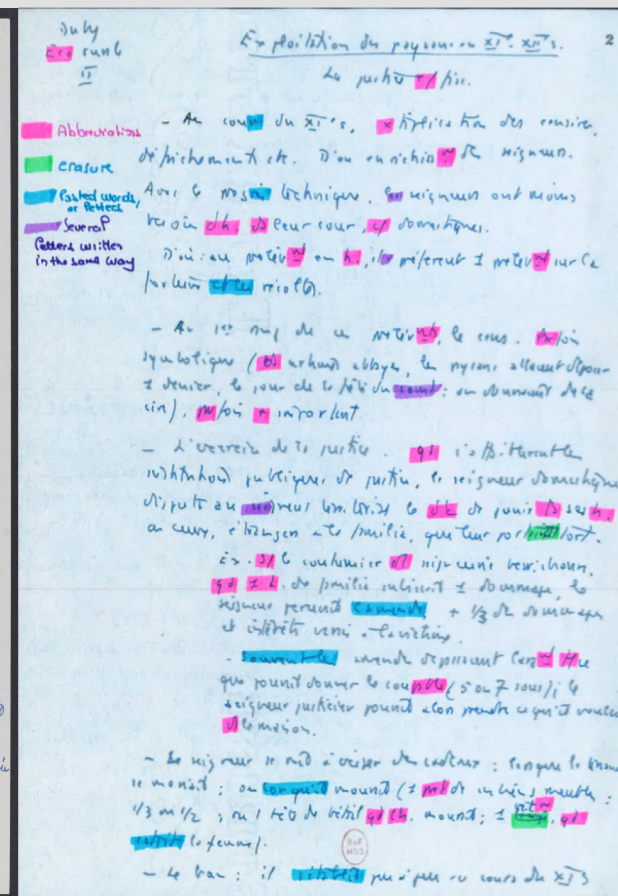
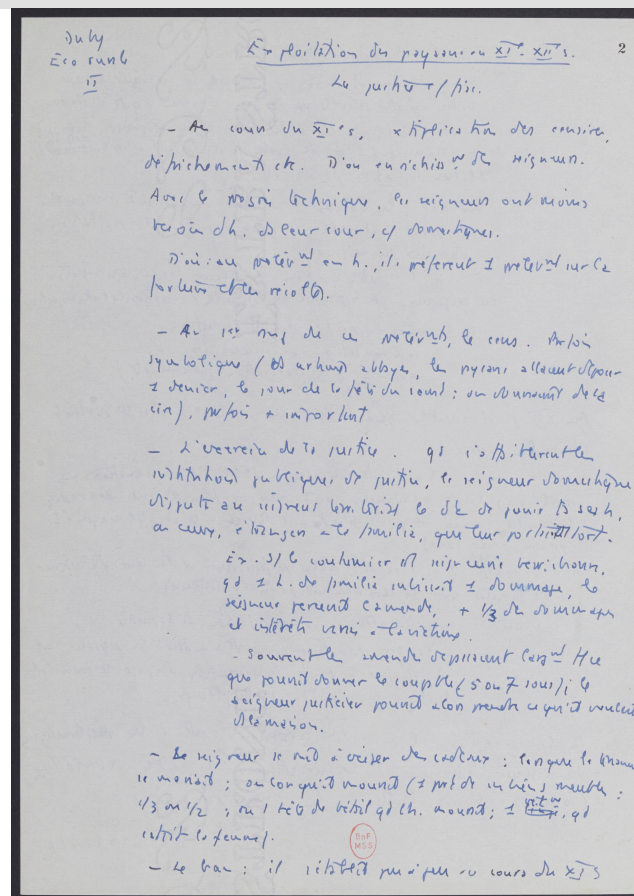
Focus sur l'expérimentation Transkribus - Créer des données d'apprentissage



Mettre en ligne, annoter et exploiter les fiches de lecture de Michel Foucault

Focus sur l'expérimentation Transkribus - Principales difficultés rencontrées

- L'écriture de Foucault : notes de lecture avec beaucoup d'abréviations ambiguës, plusieurs lettres écrites de la même manière, ratures, mots mal orthographiés, collés, etc.
- Les mots illisibles ont été identifiés grâce à l'encodage TEI.
- Prise en main du logiciel et temps moyen par carte
- Ergonomie du logiciel
- Transcriptions à plusieurs mains



Mettre en ligne, annoter et exploiter les fiches de lecture de Michel Foucault

Focus sur l'expérimentation Transkribus – Transcription automatique avec le modèle HTR

Loaded doc: boites1_diff, ID: 35689, Page 51, file: T0000424.jpeg [Image Meta Info: (Resolution:400.0, w*h: 3496 * 4916)] [current line: w*h: 2394 * 252]

Search current document...

166 - Faute des parents qui punissent les enfants

167 - "L'on ne doit presque jamais battre les enfants, car, sans compter que c'est les avilir et le ravaler au rang des malheureux, et qu'on leur inspire des sentiments bas, aimants, suivent le mensonge et pa. des vices encore + gdtis, c'est qu'il est très évident que cela est

5 des enfants

6 1762.

7 - Faute des parents qui punissent les enfants

8 166

9 lorsqu'ils se sont blessés ds des jeux, si bien que

10 les enfants cochent leurs blessures, qui denentrent

11 + grave.

12 167

13 - "L'on ne doit presque jamais battre les

14 enfants, car, sans compter que c'est les avilir et

15 le ravaler au rang des malheureux, et qu'on

16 leur inspire des sentiments bas, aimants,

17 suivent le mensonge et pa. des vices encore +

18 gdtis, c'est qu'il est très évident que cela est

Loaded doc: boites1_diff, ID: 35689, Page 45, file: T0000418.jpeg [Image Meta Info: (Resolution:400.0, w*h: 3524 * 4936)] [current line: w*h: 2299 * 224]

Search current document...

1776

122

11 L'aliment qui convient le mieux aux enfants

12 en ga jusqu'a ce que leurs forces soient accrues

13 eu x de pouvoir digérer des aliments solides est

14 iconestibint le lait et surtt celui de la mère

15 si elle est saine Le lait des mammelles de la

16 mère, qui n'est qu'une gradation du lait utérain

17 123

18 est la nourriture la + analogue la constitution

19 parhieuure de l'enfant. La nature, qui proportionne

20 tjrs les ressources au besom ... donne à chaque

1 anit

2 Essai sur la sant et

3 Le meilleur alement pr lanfant

4 léducatim medat des filles

5 le lait de la mère

6 17 76

7 122

8 11 L'aliment qui convient le mieux aux enfants

9 en ga jusqu'a ce que leurs forces soient accrues

10 eu x de pouvoir digérer des aliments solides est

11 iconestibint le lait et surtt celui de la mère

12 si elle est saine Le lait des mammelles de la

13 mère, qui n'est qu'une gradation du lait utérain

14 123

15 est la nourriture la + analogue la constitution

16 parhieuure de l'enfant. La nature, qui proportionne

17 tjrs les ressources au besom ... donne à chaque

Mettre en ligne, annoter et exploiter les fiches de lecture de Michel Foucault

Focus sur l'expérimentation Transkribus - La fonction de repérage de mots-clés (Keyword Spotting)

Documents | Fulltext (Solr) | Tags | **KWS**

Search in: ☐ Partial Matches ☒ Case-sensitivity ☐ Expert Syntax

Confidence Threshold:

Queries

Keyword 1:

Keyword 2:

Keyword 3:

Keyword Spotting Results

Confidence	Page Nr.	Line transcription	Preview
0.8061	677	D où des conflits entre le juuticier et le seigneur	
0.8052	618	Cette justierroftais au seigneur : d où le fait qd il	
0.7784	610	à son seigneur se developpèrent lavent tout en	
0.7641	605	fait simplint appe lacu seigneur territorial pr les cs	
0.7583	638	de filce le seigneur le + puissant retenant les cas qui li	
0.7528	457	. Le procureur du roi cs ou du seigneur est tprs partie	
0.7478	671	arre du pouvoir central) : mais par le seigneur	
0.7476	439	justicier n est fonde dêtre seigneur féodae de tt ce qui	
0.7410	389	A bolition du "hériot" (dt pr les seigneus de semaparer	
0.7405	629	de clase. reste que te justice est possessiond1seigneur	
0.7404	683	le seigneur dt depdait le condamne	
0.7339	679	oes Normandié tte frande sur les dts du seigneur.	
0.7304	619	d1 seigneur ; mais de personnalités diversn. Aisi en 1074	

Search Results

Created: 21.09.20 21:18:27 Com
03.04.19 13:33:08 Com
03.04.19 13:30:32 Com

Preview

- Nouvel outil puissant (calcul intensif) pour **rechercher des mots distincts** dans une collection de documents ;
- **Lancer un modèle HTR** pour produire une transcription, puis rechercher immédiatement dans le document ;
- Même si la transcription générée contient des erreurs, **KWS trouvera de manière fiable des mots, phrases, etc.** dans les documents ;
- Le programme indique le nombre d'occurrences trouvées, les pages et les images et évalue la fiabilité des résultats ;
- Dans notre ex., le programme retrouve 259 fois « justice » et 154 fois « seigneur » (bien + qu'un moteur de recherche classique). On remarque que même si la transcription est fautive « Ceigneur » (p. 457), le KWS a bien identifié l'image du mot.

Boite_001 | Système pénal. Moyen-âge, XVIe siècle.

Mettre en ligne, annoter et exploiter les fiches de lecture de Michel Foucault

Focus sur l'expérimentation Transkribus - Résultats et perspectives

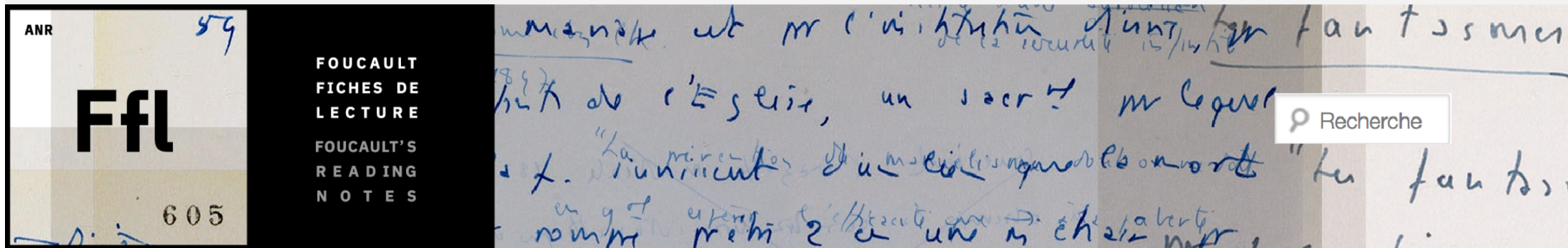
- Entraînement du modèle HTR sur **600 images** transcrites **manuellement**, puis des transcriptions automatiques corrigées : **95% des caractères reconnus** sans erreur
- De nouvelles transcriptions automatiques peuvent être générées et **corrigées** et de **nouveaux modèles peuvent être entraînés** pour améliorer encore les résultats
- Les **résultats des tests** du logiciel sont très **encourageants** : Certes, il reste des erreurs... mais avec un taux d'erreur de 5% au caractère, les fiches deviennent **lisibles et indexables** par un moteur d'indexation*.
- Produire des **transcriptions automatiques pour toutes les images** et **indexation** du corpus (ouvrages, auteurs, concepts, etc.)
- **Importer les transcriptions automatiques dans FFL-Eman (Omeka)** – (script PHP développé par V. Ventresque)
- Utiliser la plateforme FFL-Eman pour mettre en place un **système de correction collaborative** ou l'interface Web proposée par Transkribus
- **Keywords spotting (KWS)** cad recherche approximative sur l'image : un véritable + même dans le cas de transcription fautive / facilite l'exploration

*L'**indexation automatique de documents** est un domaine de l'informatique et des sciences de l'information et des bibliothèques qui utilise des méthodes logicielles pour organiser un ensemble de documents et faciliter ultérieurement la recherche de contenu dans cette collection.



Malgré l'imperfection des transcriptions, le dispositif s'avère extrêmement utile pour l'exploration d'un corpus. Tester l'outil expert Transkribus a modifié en profondeur notre point de vue sur le corpus et son exploration.

<https://ffl.hypotheses.org/>



RETROUVEZ L'ACTUALITÉ DU PROJET FFL : [HTTPS://FFL.HYPOTHESES.ORG/](https://ffl.hypotheses.org/)

ET LES VIDÉOS DU COLLOQUE FINAL À A BNF : [HTTPS://FFL.HYPOTHESES.ORG/2092](https://ffl.hypotheses.org/2092)