

# ÉTUDIER LE « GROUPE DE NEUCHÂTEL » : DE L'ÉDITION DES *FAITS* À UN CORPUS NUMÉRIQUE DE LA PREMIÈRE RÉFORME ROMANDE

SONIA SOLFRINI<sup>1</sup>, GENEVIÈVE GROSS<sup>1</sup>, BRIGITTE ROUX<sup>1</sup>, NATHALIE SZCZECH<sup>2</sup>,  
PIERRE-OLIVIER BEAULNES<sup>1</sup>, AURÉLIA M. OLIVEIRA<sup>1</sup>, ET DANIELA SOLFAROLI CAMILLOCCI<sup>1</sup>

<sup>1</sup>Institut d'Histoire de la Réformation, Université de Genève (Suisse)

<sup>2</sup>Département d'histoire, Université Bordeaux-Montaigne (France)

## INTRODUCTION

La constitution d'un corpus numérique des textes de la première Réforme romande s'inscrit dans le cadre du projet SETAF. Une importante dimension numérique est prévue pour la réalisation de ce projet, avec la numérisation de nombre d'imprimés d'une part (segmentation, OCR et encodage TEI), et l'exploitation computationnelle de ces données d'autre part (analyse linguistique et stylométrique).

## LE PROJET SETAF

Le projet SETAF (S'En Tenir Aux *Faits*) entend mener une ample investigation historique à caractère pluridisciplinaire à partir d'un ouvrage singulier émanant du milieu éditorial du « groupe de Neuchâtel », réuni autour du réformateur Guillaume Farel (1489-1565). L'ouvrage autour duquel s'articule le projet sont les *Faits de Jésus Christ et du pape* : paru sous couvert d'anonymat dans trois éditions successives, entre les années 1530 et 1560, il s'agit de l'unique livre illustré de la Réforme francophone conservé pour la première moitié du XVI<sup>e</sup> siècle. L'équipe fournira, sous forme numérique, l'édition critique des *Faits* et un catalogue raisonné de ses images. En parallèle, l'équipe construira un corpus numérique comprenant les imprimés évangéliques romands, notamment les ouvrages écrits et réédités par Farel et son cercle à partir des années 1530. Jusqu'ici, nous avons établi un corpus d'une cinquantaine d'ouvrages, publiés par les imprimeurs Pierre de Vingle et Jean Michel, à Neuchâtel et à Genève, auxquels seront associés d'autres imprimés relevant du même réseau.

## LE CORPUS NUMÉRIQUE

L'océrisation des textes du corpus est produite par l'équipe grâce à l'instance genevoise FoNDUE [2] d'*eScriptorium* [3] et aux modèles du projet *Gallicorpora* comme *Cortado* [4]. Deux types de modèles sont actuellement entraînés : un premier pour la segmentation (cf. fig. 1), et un second pour la reconnaissance optique de caractères (cf. fig. 2). L'encodage des textes en XML-TEI suivra autant que possible les recommandations des BVH (*Bibliothèques Virtuelles Humanistes*) du CESR [5] et celles du projet *Gallicorpora* [6]. Il est également envisagé d'annoter linguistiquement les textes grâce aux modèles du projet *FreEM* [7] afin de simplifier l'étude des données. Ce corpus permettra d'interroger l'intertextualité à l'œuvre dans les *Faits* et servira à l'étude de questions linguistiques et stylistiques. Au niveau sémantique, par exemple, nous pensons recourir à la sémantique distributionnelle ; au niveau stylométrique, plusieurs textes de ce corpus étant anonymes, la question de leur attribution pourrait être abordée sous un nouvel angle.

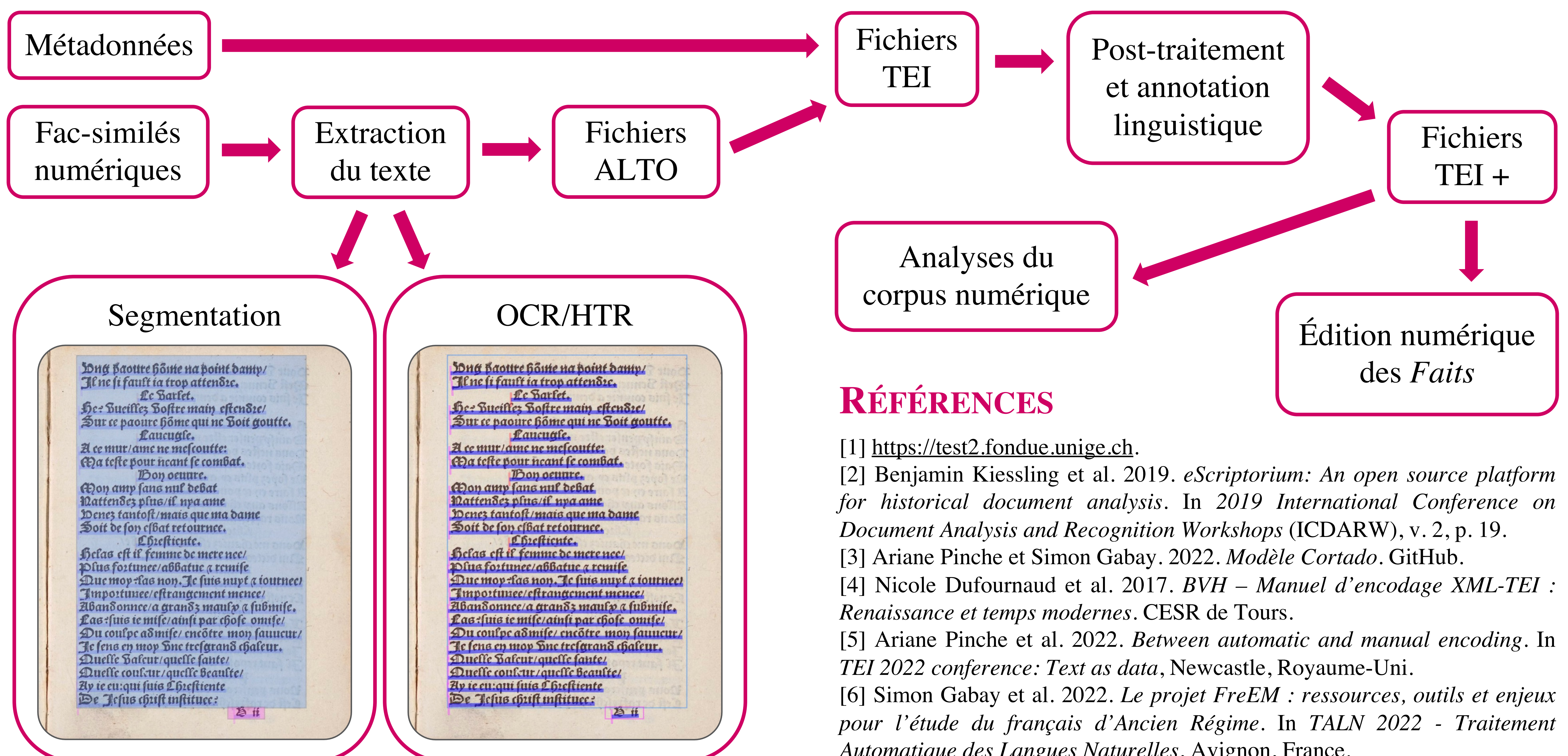


Fig. 1. Vérité terrain pour l'entraînement du modèle de segmentation [7].

Fig. 2. Vérité terrain pour l'entraînement du modèle d'OCR [7].

## RÉFÉRENCES

- [1] <https://test2.fondue.unige.ch>.
- [2] Benjamin Kiessling et al. 2019. *eScriptorium: An open source platform for historical document analysis*. In *2019 International Conference on Document Analysis and Recognition Workshops (ICDARW)*, v. 2, p. 19.
- [3] Ariane Pinche et Simon Gabay. 2022. *Modèle Cortado*. GitHub.
- [4] Nicole Dufournaud et al. 2017. *BVH – Manuel d'encodage XML-TEI : Renaissance et temps modernes*. CESR de Tours.
- [5] Ariane Pinche et al. 2022. *Between automatic and manual encoding*. In *TEI 2022 conference: Text as data*, Newcastle, Royaume-Uni.
- [6] Simon Gabay et al. 2022. *Le projet FreEM : ressources, outils et enjeux pour l'étude du français d'Ancien Régime*. In *TALN 2022 - Traitement Automatique des Langues Naturelles*, Avignon, France.
- [7] Matthieu Malingre. 1533. *Moralité de la maladie de Chrétienté*. Pierre de Vingle : Neuchâtel, Suisse. [10.3931/e-rara-563](https://nbn-resolving.org/urn:nbn:ch:gn-d-10.3931/e-rara-563).