

Crimes et châtements : un projet pionnier pour les procès criminels

Avec Crimes et châtements, la Fondation des Archives de l'ancien Évêché de Bâle (AAEB, Porrentruy) développe un concept innovant pour mettre en valeur les sources criminelles d'Ancien Régime tout en renouvelant les formes de consultation et de médiation des archives.

Des sources extraordinaires

Vous cherchez des informations sur la sorcellerie, les pratiques magiques sous l'Ancien Régime, les méthodes de guérison, le « secret », la sexualité, le sort des « filles-mères », les yéniches, la criminalité, la justice, la torture, le patois, l'évolution de l'orthographe, les *fake news* d'autrefois, le mobilier des paysans, les récits de vie, les famines... ? Les archives criminelles sont pour vous, car elles offrent une source inépuisable d'informations sur la vie quotidienne, les croyances et les rapports sociaux dans les communautés d'Ancien Régime. Les AAEB conservent une documentation très riche en la matière : [leurs procès criminels et de sorcellerie vont de 1461 à 1797](#) et représentent environ 160'000 pages de documents couvrant plus de 25 mètres d'étagères.

***Crimes et châtements* : le concept**

Faute de répertoire, cette belle documentation restait toutefois inexploitée. Les premiers travaux d'inventorisation ont confirmé sa richesse, mais mis en évidence deux problèmes. D'abord, le grand désordre de ces séries impliquait un traitement pièce à pièce ; cela nous a conduits à opter pour des répertoires extrêmement détaillés. Ensuite, ces manuscrits sont souvent difficiles à lire, y compris pour les académiques. En effet, ils posent d'importants problèmes paléographiques, d'autant que les dossiers comprennent fréquemment des documents en allemand, rédigés en *Kurrentschrift*. Lancé à la fin de 2021, le projet *Crimes et châtiments* offrira au public le meilleur accès possible à ces sources grâce à des outils complémentaires consultables en ligne :

- Des répertoires extrêmement détaillés (ex. en français : homicide « intercantonal » en 1616, [PCrim E 78](#) ; ex. en allemand : Sodomie mit Pferden, 1609, [PCrim LZ 25](#)).
- Les reproductions numériques des documents et, lorsqu'elles sont déjà disponibles, leurs transcriptions générées automatiquement (HTR : *Handwritten Text Recognition*) et progressivement corrigées (ex. : interrogatoire d'une femme soupçonnée de sorcellerie avec le texte d'une prière de guérison, 1596, [B 168/14-37.2](#), p. 1)

La rédaction des répertoires, la mise en ligne des reproductions et l'établissement des liens entre les sites seront achevés en 2027. La création et la correction des transcriptions automatiques prendront naturellement plus de temps et se poursuivront par la suite.

Reproductions et stockage des images

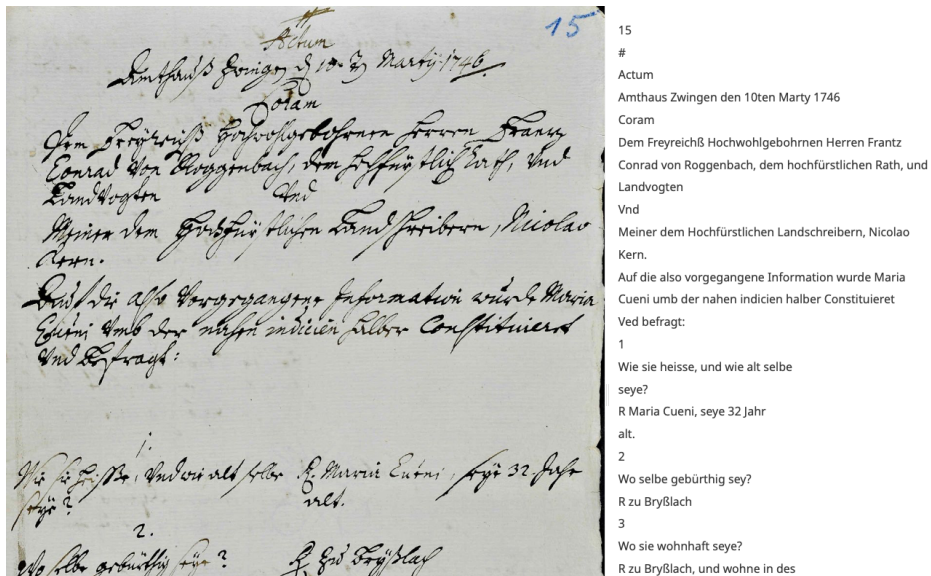
Les documents d'archives sont photographiés selon les standards suivants :

- Numérisation en couleur (format en 24 bits).
- Résolution de 300 DPI, permettant une restitution parfaite à la même taille que le document original.
- Formats TIFF pour les fichiers maîtres et JPEG pour les fichiers dérivés destinés à être mis en ligne.
- Production de MD5 pour chaque image, lesquels permettent de vérifier l'intégrité des données.

L'archivage sécurisé et à long terme des documents numérisés est géré par le Service de l'informatique du Canton du Jura (SDI), dans le cadre du portefeuille de projets de numérisation SIGMA de l'Office de la culture. Les fichiers TIFF sont livrés au SDI avec les MD5 correspondants, garants de l'intégrité des données. Les fichiers numériques sont stockés à long terme sur des serveurs externes et sont quotidiennement sauvegardés dans des *Datacenters* situés dans des localités différentes.

Transcriptions automatiques (HTR) en allemand et en français

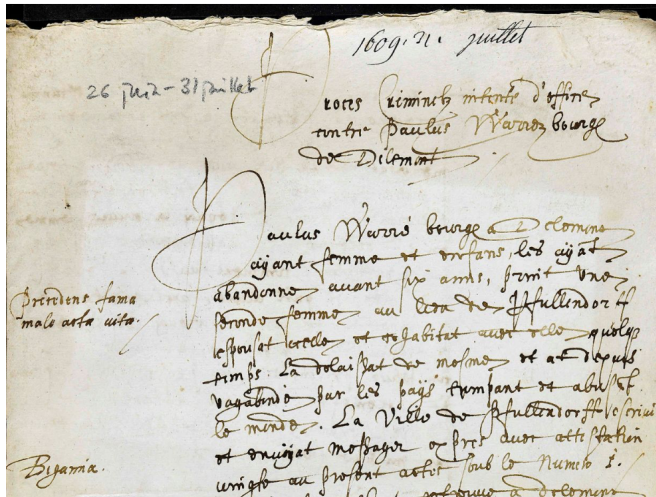
Transcription en allemand : Transkribus



Interrogatoire de Maria Cueni, de Brislach, soupçonnée d'inceste avec son frère (1746, PCrim LZ 138-3, haut de la p. 1).

Le point crucial et le plus novateur du projet est de générer des transcriptions automatiques. Dans ce but, les AAEB ont adhéré à la coopérative [Transkribus](#), la plus grande plateforme de reconnaissance de texte et d'analyse d'images de documents historiques, issue de deux programmes de recherche européens coordonnés par l'université d'Innsbruck. Le choix de Transkribus s'est imposé, car cette plateforme offre des services à un prix abordable et dispose de modèles de transcription très efficaces pour les textes en allemand. D'autres plateformes existent, en particulier [FoNDUE](#) (Formes Numérisées et Détection Unifiée des Écritures), une application d'[e-scriptorium](#) développée à l'université de Genève, avec laquelle nous collaborons en partageant nos données. La possibilité d'utiliser l'infrastructure FoNDUE a été évaluée en début de projet. La solution tout intégrée de Transkribus, du développement des modèles à la publication des données, a été préférée puisqu'elle nécessite moins de compétences informatiques.

Transcriptions des textes en français : l'apport majeur des AAEB



1609. 31 juillet
 26 juin - 31 juillet
 Procès criminel intenté d'office
 contre Paulus Warré bourgeois
 de Delemont,
 Precedens fama
 male acta vita.
 Paulus Warré bourgeois a Delemont
 ayant femme et enfans, les ayant
 abandonné avant six ans, print une
 seconde femme au lieu de Pfullendorff
 espousa icelle et cohabitait avec elle quelque
 temps, la delaisa de mesme et at depuis
 vagabondé sur les pays trompant et abusant
 le monde. La ville de Pfullendorff escrivit
 et envoya messenger expres avec attestation
 jointe au present actes soub le Numero 1
 Bigamia
 Ledit Paulus cestant retreuve a delemont
 a la St Martin passé et advertit que son
 second mariage estoit descovert craignant
 estre saysy prisonnier soy retirat apres
 Fuga

Rapport sur l'arrestation de Paul Varré, de Delémont (maître d'école, soldat et...
 pèlerin) accusé de trigamie, vols, escroqueries et violences (1609, PCrim Dt
 (ville) 6-1, p. 1)

En l'absence de modèle satisfaisant pour les textes en français, les AAEB ont dû créer les leurs, au prix d'un très grand travail – toujours en cours. La méthodologie a consisté à spécialiser sur nos données un modèle public développé par le prof. Tobias Hodel (*fine tuning*). Dans une première étape, l'algorithme a été exercé pour déchiffrer l'écriture d'un fonctionnaire du début du 17^e siècle, le prévôt Henri Farine, qui a rédigé une bonne partie des documents liés à la chasse aux sorcières dans les années 1600-1618. Actuellement, ce modèle donne d'excellents résultats (*Character Error Rate* ou CER de 4.4 %). Sur cette base, nous avons développé un modèle mixte en ajoutant des textes de trois autres officiers bien représentés dans les sources et dont les écritures sont très différentes de celle de Farine. Les résultats encourageants obtenus nous ont permis de diversifier le modèle en intégrant aux données d'entraînement et d'évaluation des transcriptions supplémentaires de « mains » diverses du 16^e et du 17^e siècle : le dernier item affiche un CER de 8.2 %, avec de belles perspectives d'amélioration, étant entendu que ses performances varient fortement selon les écritures. Notre modèle sera bientôt rendu public sur le site de Transkribus et pourra servir à d'autres institutions. Il faut souligner que les textes du 16^e et du 17^e siècle utilisés par le modèle des AAEB constituent actuellement le corpus le plus important pour des textes en français de cette période !

Communication et médiation

Crimes et châtements suscite un grand intérêt dans les médias, qui ont largement rendu compte des résultats intermédiaires présentés lors de notre conférence de presse du 28 août 2024 et ont parfois consacré de pleines pages ou des séries à ce sujet (voir [revue de presse](#)). Par ailleurs, la cheffe de projet, Élodie Paupe, et le conservateur, Jean-Claude Rebetez, participent à des colloques et des tables rondes, rédigent des articles ou donnent des conférences qui visent à faire connaître *Crimes et châtements* par les professionnel.le.s. et le grand public. À l'intention de ce dernier, le site web des AAEB a été enrichi d'une introduction sur ce sujet, mettant particulièrement en valeur la [chasse aux sorcières](#) avec une présentation générale du sujet, un « Abécédaire » de la sorcellerie, une liste nominative de toutes les personnes ayant fait l'objet d'un procès ou ayant été dénoncées comme suspectes de sorcellerie. D'autres thèmes seront bientôt présentés sur le site, comme la condition des femmes ou les formes de la torture judiciaire.

Visions d'avenir et *crowdsourcing*

Crimes et châtements permettra une évolution des pratiques de consultation des lecteurs et des lectrices : sur la base de l'inventaire en ligne, ils et elles accèderont non seulement à la reproduction du document décrit, mais aussi à sa transcription, si elle existe déjà. Dans le cas contraire, ils et elles pourront demander qu'une transcription automatique soit générée par les archivistes. D'ici la fin du projet, les AAEB développeront un concept pour permettre au public de contribuer à la correction de ces nouvelles transcriptions, qui seront dès lors accessibles à tous et toutes. Il s'agit d'une innovation majeure qui bouleversera la consultation des archives et apportera un gain scientifique durable.

Dans le deuxième semestre 2025, les AAEB organiseront un atelier ouvert aux collègues archivistes et bibliothécaires intéressé.e.s par notre projet et le HTR. La date et les informations seront communiquées sur le site d'*arbido*.



Jean-Claude Rebetez

Jean-Claude Rebetez est historien, archiviste-paléographe et conservateur des Archives de l'ancien Evêché de Bâle. Il donne le cours de paléographie médiévale et du 16^e siècle à l'université de Bâle.

Abstract

Deutsch

Das Projekt *Crimes et châtiments* hat sich zum Ziel gesetzt, alle Kriminal- und Hexenprozesse aus dem *Archives de l'ancien Évêché de Bâle (AAEB)* vom Ende des 15. Jahrhunderts bis zur Französischen Revolution online zu stellen. Sehr detaillierte Verzeichnisse ermöglichen es, die Dokumente zu identifizieren, während ein Link zu ihrer digitalen Reproduktion und (sofern vorhanden) Transkription führt. Verschiedene Vermittlungsmethoden erleichtern den Benutzenden die Orientierung auf der Plattform. Die grosse Neuheit des Projekts besteht im Online-Zugang zu automatisch korrigierten Transkriptionen (ein *Crowdsourcing* Konzept zur Transkription-Korrektur wird bald angeboten). Das AAEB spielt damit eine Vorreiterrolle in der Entwicklung eines Modells für automatische Transkriptionen für französische Texte aus dem 16. und 17. Jahrhundert.

Français

Crimes et châtiments a pour but de mettre en ligne l'ensemble des procès criminels et de sorcellerie des Archives de l'ancien Évêché de Bâle (AAEB) de la fin du 15^e siècle à la Révolution française. Des répertoires très détaillés permettent d'identifier les documents et un lien mène à leur reproduction numérique et à leur transcription (quand elle existe). Divers moyens de médiation facilitent l'orientation du public. La grande originalité du projet consiste dans la mise en ligne de transcriptions automatiques corrigées (un concept de travail participatif pour ces corrections sera bientôt proposé au public). Les AAEB jouent un rôle pionnier dans l'élaboration d'un modèle de transcription automatique pour les textes en français des 16^e et 17^e siècle.