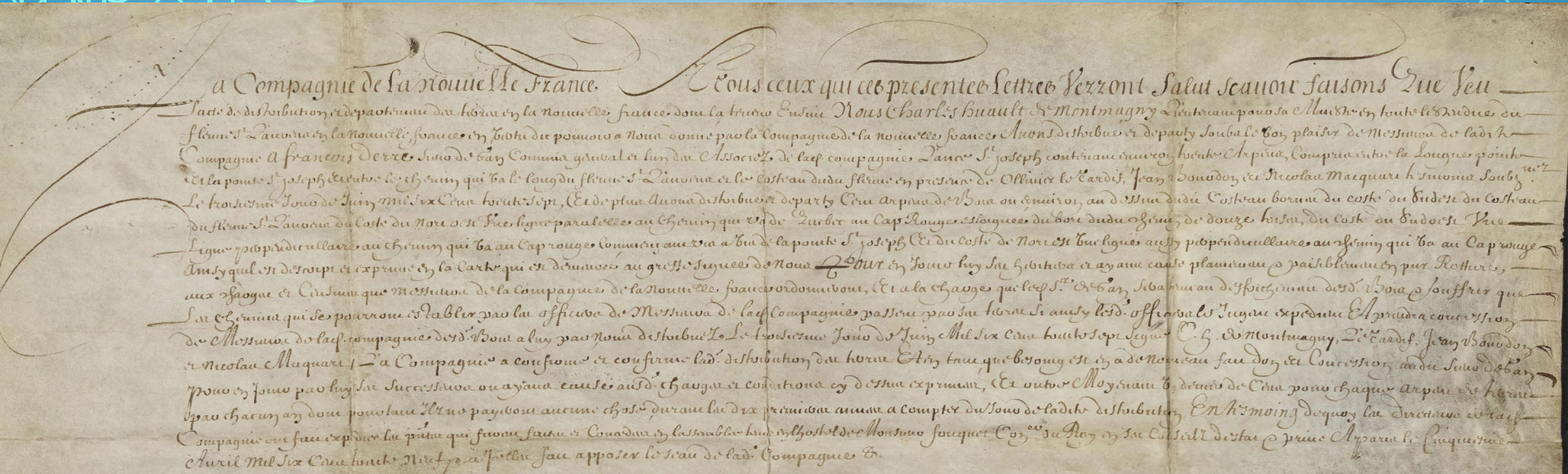


Nouvelle-France numérique

Une expérience de partenariat global dans la recherche sur le patrimoine documentaire et la gestion des données



Maxime Gohier

Université du Québec à Rimouski

Collections culturelles et patrimoniales numériques de l'avenir : élargir la recherche et favoriser un accès novateur

Conseil des chercheurs du RCDR

Montréal

2 novembre 2022



Nouvelle-France
NUMÉRIQUE

NOUVELLE- FRANCE NUMÉRIQUE

Objectifs du projet :

- Scientifique :
 - Étudier la production et la circulation de l'information à l'époque de la Nouvelle-France
- Méthodologique :
 - Repérer, dans de très vastes corpus documentaires, des indices d'intertextualité (emprunts textuels, structures stylistiques récurrentes, marques de modifications)
 - Identifier les acteurs et les réseaux qui les unissent
- Technologique :
 - Identifier les technologies permettant à tous les chercheurs sur la Nouvelle-France de mettre en commun leurs données de recherche
 - Faciliter l'analyse de ces données avec de nouveaux outils technologiques

Caroles des proquois qui
ont demandé de leur pain à Montréal
au Sr. P. L. Bruyart et les^{rs} de
Mairourses de joncuire qui ont
amené des depuzes de la nation
au nombre de dix neuf pour la conclusion
de la paix, jls ont dit ce qui suit à
Monsieur le Chevalier de Callieres
Gouverneur en Lieutenant Général pour
le Roy en toute la France Septentrionale
le 3. d'Avr. 1700

Par M^{re} Corde de porcelaine
Pour Ceste fois Notre pere omniscient
vous voyez jay touché les nations
proquoises c'est vray que vous n'y voyez
par le visage de bonneyous parce que
celuy qui étoit depuzé pour venir, est

NOUVELLE-FRANCE NUMÉRIQUE

- Partenariat de recherche (CRSH)
- Une vingtaine de chercheurs de différentes disciplines (histoire, lettres, géographie, anthropologie, linguistique, histoire de l'art, études en patrimoine, archivistique, etc.)
- Partenaires institutionnels :
 - BAnQ
 - BAC
 - Musée de la Civilisation
 - Archives et Bibliothèque des livres rares, Université de Montréal
 - READ-COOP SCE (*Transkribus*)
- Citoyens bénévoles

TRANSKRIBUS : L'IA AU SERVICE DE L'ANALYSE DES MANUSCRITS ET DE LA GESTION DES DONNÉES DE RECHERCHE

Transkribus Expert Client v1.21.0 (22_09_2022_07:35), Loaded doc: ANOM, C13A, vol. 1, ID: 496877, Page 39, file: FRANOM34_C13_001_0039.jpg [Image Meta Info: (Resolution:200.0, w*h: 2248 * 3152)] [current l...

Search current document... 39 /682 Final

Server Overview Layout Metadata Tools

Logout maxime_gohier@uqar.ca

Document... Find

Document Manager User Manager

Versions Jobs

Recent Documents... User activity

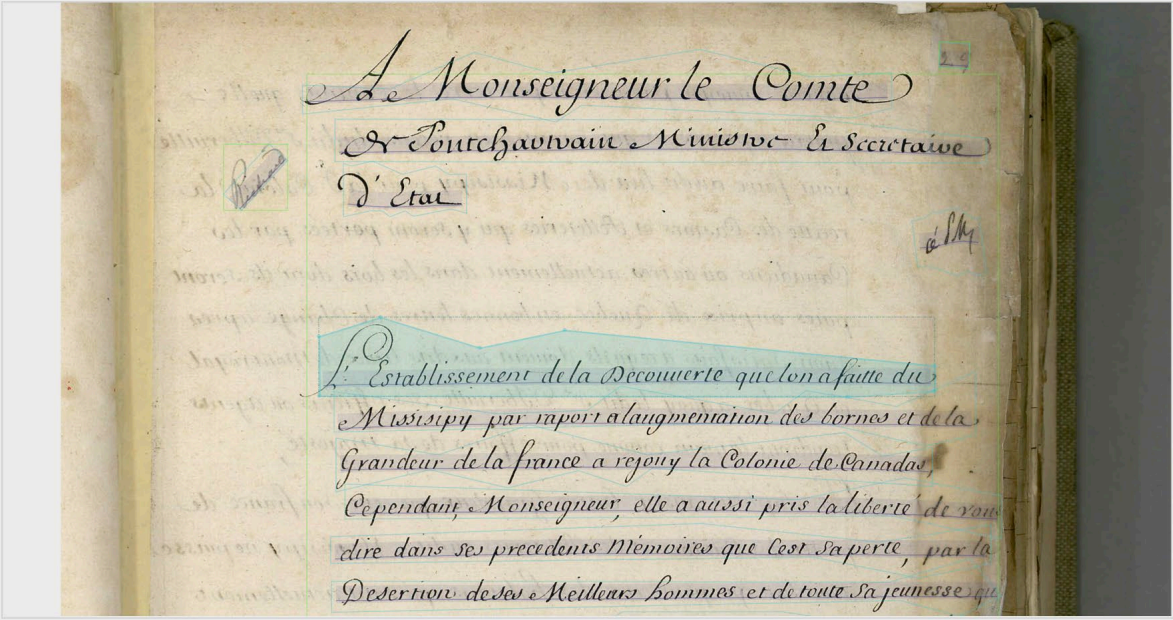
Collections: ANOM, C13A (55686, Owner) Col-ID

Documents Model Data

1-54 / 54 1 1

ID	Title	Pages	Uplo
49689E	ANOM, C13A, vol. 41	1010	max
49689E	ANOM, C13A, vol. 40	798	max
496894	ANOM, C13A, vol. 9	876	max
496893	ANOM, C13A, vol. 8	956	max
496892	ANOM, C13A, vol. 7	666	max
496891	ANOM, C13A, vol. 6	894	max
49689C	ANOM, C13A, vol. 5	762	max
49688E	ANOM, C13A, vol. 4	1094	max
49688E	ANOM, C13A, vol. 3	970	max
496887	ANOM, C13A, vol. 2	900	max
49688E	ANOM, C13A, vol. 10	740	max
496877	ANOM, C13A, vol. 1	682	max

100 Filter



3-1 A Monseigneur le Comte

3-2 de Pontchartrain Ministre Et Secretaire

3-3 d'Etat

3-4 à PM

3-5 L'Establissement de la Decouverte que l'on a faite du

3-6 Missisipy par raport a l'augmentation des bornes et de la

3-7 Grandeur de la france a rejouy la Colonie de Canadas,

FORMATION D'UNE COMMUNAUTÉ GLOBALE

Protocoles de travail

- Transcription standardisée
- Schéma de balisage TEI
- Outils de gestion des projets et de l'avancement des transcriptions

Développement de modèles de HTR

- Modèles privés
- Modèles publics (données ouvertes)

Formations à *Transkribus*

- Écoles d'été
- *Atelier permanent d'analyse documentaire*
- Webinaires
- Formations *ad hoc*

Les Gardenotes

- Structuration d'une équipe de transcription participative

Préparation d'un dépôt numérique

- Transcrire la Nouvelle-France

c. anbedatsin. avoir froid
 c. anbi. bande. ruban.
 c. anbidé. dents.
 c. anbin. vanbin. voir.
 c. anbesépin. juer.
 c. anbi. bouillon.
 e. anbi. moins. peu. petitement.

Archives du Séminaire d



Carte d'une Partie du Lac

Supérieur avec la Découverte de la Rivière depuis le Grand Portage. A
 jusqu'au Lac Quimpigon, ou on a Construit Le Fort Maurepas, Le Fort St.
 Charles, est dans le Lac des Bois, et dans le Lac Tekamamion, un troisième
 Fort qui porte le nom de St. Pierre.

Les Lacs et Rivière Marquette en Couleur Bleu, sont celles qu'on a découvertes

RÉSULTATS

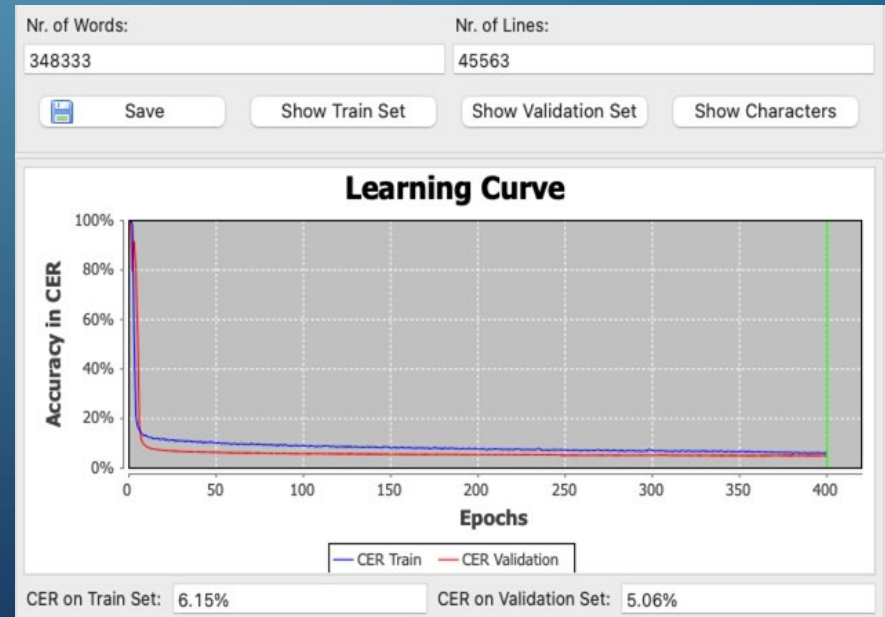
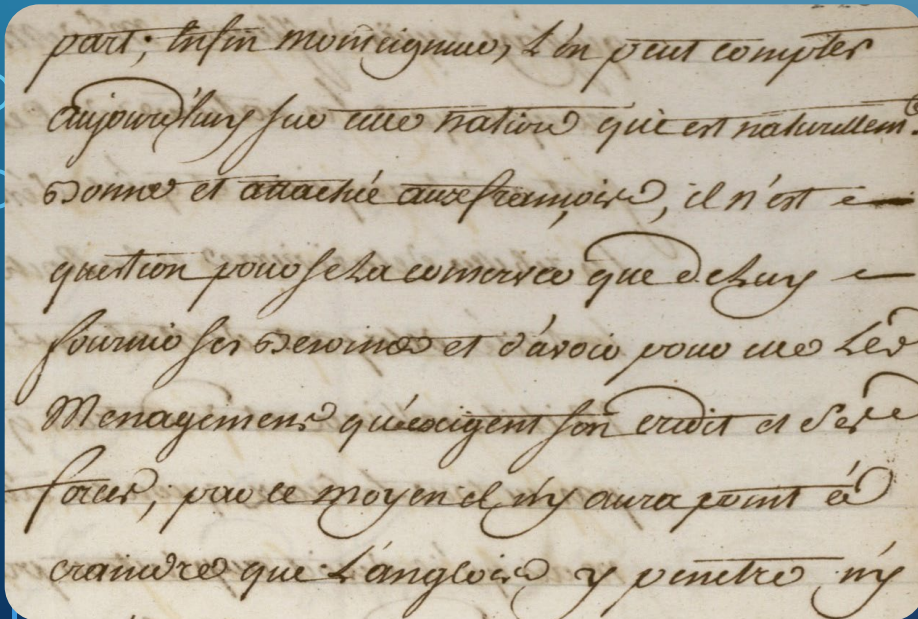
- Plus de 150 000 pages transcrites
- Deux modèles publics
- Des dizaines de modèles privés
- Grande variété de documents traités :
 - Archives judiciaires : baillage de Montréal
 - Correspondances d'administrateurs coloniaux : Louisiane et Canada
 - Cahiers d'ordonnances d'intendants
 - Greffes de notaires
 - Dictionnaires de langues autochtones
 - Cartes manuscrites

MODÈLES PUBLICS : LES CORRESPONDANCES GÉNÉRALES

- Variabilité des mains d'écriture
- Écritures de secrétaires et d'administrateurs
 - Graphies rondes, bâtardes et coulées
- Taux d'erreurs réel sur la page ci-dessous :
CER : 4,91% WER : 15,15%

Une dizaine de fonds d'archives sur un siècle (~1660-1760)

Données d'entraînement : **1832 pages**
Données de validation : **197 pages**
Total: **2029 pages**

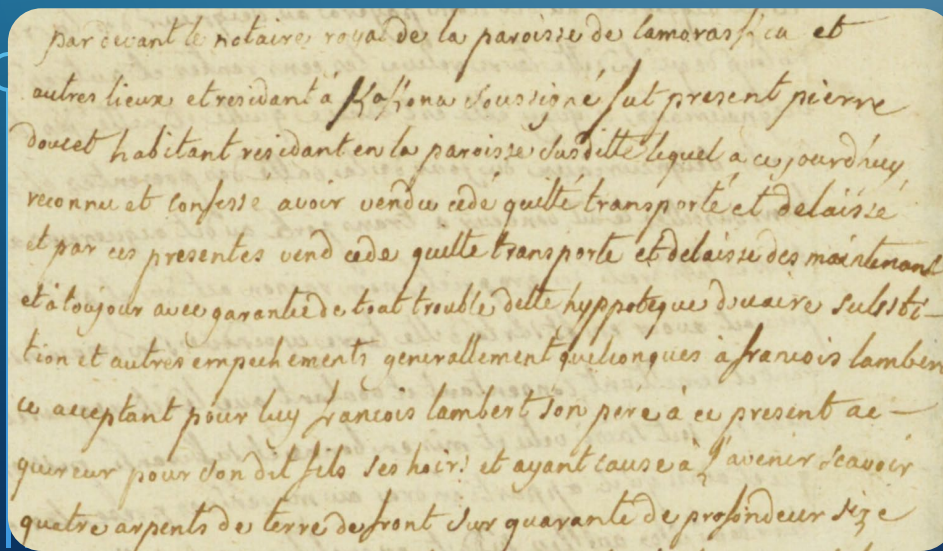


MODÈLES DÉVELOPPÉS : NOTAIRES DES 17^E ET 18^E SIÈCLES

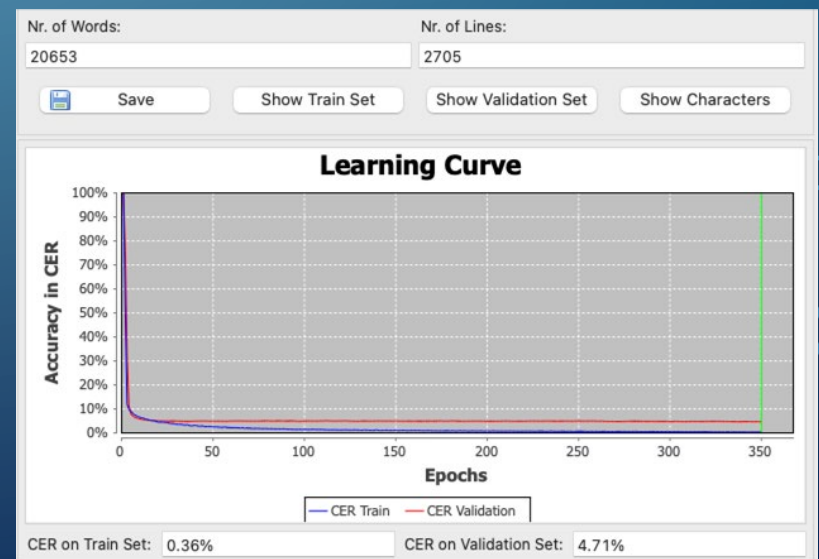
- Régularité des mains d'écritures
 - Un greffe = un notaire
- Lexique standardisé : termes juridiques récurrents
- Taux d'erreurs réel sur la page ci-dessous :
CER : 1,35% **WER : 7,64%**
- Qualité de l'écriture très variable selon les notaires
- Modèle très performant avec peu de données d'entraînement

Greffe de Michel Saindon (1768-1780)
BAnQ-Rimouski, CN104,S50

Données d'entraînement : **112 pages**
Données de validation : **27 pages**
Total : **139 pages**



par devant le notaire royal de la paroisse de Lamoras, et
autres lieux et résidant à Houlton, sous signé fut présent pierre
Doucet habitant résidant en la paroisse du dit lieu lequel à ce jourd'hui
reconnu et confesse avoir vendu cédé quitte transporté et delaisié
et par ces présentes vend cédé quitte transporté et delaisié des maintenant
et à toujours avec garantie de tout trouble de cette hypothèque de vaux subsisti-
tion et autres empêchements généralement quelconques à François Lambert
ce acceptant pour lui François Lambert son père à ce present ac-
querir pour son dit fils ses heirs et ayant cause à l'avenir de avoir
quatre arpents de terre des front sur quarante de profondeur six



PROCHAINES ÉTAPES

Accroître le nombre
de partenaires

Développer le dépôt
numérique *TNF*

Développer des outils
d'IA pour le
traitement des
données

Étudier les possibilités
de liage des données
textuelles avec
d'autres types de
données

Diffuser le modèle
pour d'autres projets