

L'exploration visuelle comme instrument de recherche en histoire économique et sociale

Paul Girard, Guillaume Daudin, Loïc Charles, Guillaume Plique

▶ **To cite this version:**

Paul Girard, Guillaume Daudin, Loïc Charles, Guillaume Plique. L'exploration visuelle comme instrument de recherche en histoire économique et sociale. Humanistica 2020, May 2020, Bordeaux, France. hal-04076251

HAL Id: hal-04076251

https://sciencespo.hal.science/hal-04076251

Submitted on 20 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.

L’exploration visuelle comme instrument de recherche en histoire économique et sociale

Paul Girard (1), Guillaume Daudin (2), Loïc Charles (3), Guillaume Plique (4)

(1) OuestWare, Nantes, France; (2) Université Paris-Dauphine, PSL Research University, CNRS , UMR 8007, IRD, UMR 260, LEDa, DIAL, Paris, France; (3) Led (EA 3391), Université de Paris 8 & Institut national d’études démographiques, France; (4) Sciences Po, médialab (EA 7033), Paris, France.

Une recherche par le visuel

La démocratisation de méthodes de visualisation de données a généré un regain d’intérêt pour l’analyse statistique exploratoire . Mais les humanités ne peuvent se saisir des nouveaux outils de visualisation que s’ils permettent des interprétations multiples respectant la complexité des terrains et des questions de recherche . De meilleures interactions homme-données permettraient de dépasser un formalisme trop radical des images en offrant la possibilité d’explorer la complexité sans la réduire *a priori* . C’est dans cette perspective que nous avons développé une nouvelle famille d’instruments de recherche visuels : les “datascapes” .
Afin de présenter les possibilités apportées par l'exploration visuelle pour développer une interprétation complexe de jeux de données, nous présentons[1] un datascape permettant d'explorer le commerce de la France au XVIIIe siècle. Cet outil, un assemblage hypertextuel de plusieurs visualisations de données, permet au chercheur de construire un chemin interprétatif jalonné de différentes perspectives sur un corpus de données historiques hétérogènes. Tout au long de ce parcours, des moyens de filtrage avancés et un système de classification hiérarchique permettent d'adapter chaque visualisation à la question de recherche.

Les échanges commerciaux de la France au XVIIIe siècle

Notre cas d’étude est issu du projet ANR TOFLIT18 portant sur les échanges commerciaux de la France au XVIIIe siècle. Le projet a compilé les archives du Bureau de la Balance du Commerce (créé en 1713) et du Bureau des archives du commerce (après 1792) . Nous avons créé une base de données répertoriant les flux commerciaux décrits par ces sources : chaque flux indique des informations sur un produit échangé (désignation, prix, quantité…) entre des points d’enregistrement en France (“Directions”) et divers partenaires commerciaux de 1716 à 1821 . Nous avons construit un instrument de recherche visuel qui propose cinq outils de visualisation :

- *Time series* : séries temporelles variées.
- *Locations network* : réseaux biparties Direction-Partenaire commercial révélant la

structure territoriale du commerce de la France.

- *Product terms network* : réseaux des séquences de termes utilisés dans les noms de marchandises.
- *Metadata* : projections du nombre de points de données de la base suivant de nombreux axes pour évaluer la couverture du corpus.
- *Classifications* : recherches dans le système de classification hiérarchique utilisé pour manipuler les noms de produits et de pays transcrits des sources.

Afin d'éclairer la pertinence des dynamiques exploratoires de ce datascape, nous illustrons dans la section suivante son usage dans le cadre d'une question de recherche en histoire économique : l'étude des conséquences économiques de la perte du Canada (Nouvelle France) par la France au profit du Royaume Uni en 1763 suite à la guerre de Sept ans.

Effets économiques de la perte du Canada en 1763

De 1756 à 1763, la guerre de Sept Ans opposa les principales puissances européennes et leur possessions coloniales . Ce conflit est régulièrement cité comme le premier conflit global car des combats importants se déroulèrent loin de l'Europe. Le traité de Paris concluant le confit est une humiliation pour la France qui perd une large part de son empire colonial et notamment la Nouvelle France (possession de le France en Amérique du nord). Bien que considérable en terme de surface, cette perte est depuis le XVIIIe siècle considérée comme marginale d'un point de vue économique et par conséquent peu étudiée.
Le corpus et l'application web TOFLIT18 permettent d'étudier l'impact de la perte du Canada sur l'économie française tant au niveau national que local. En mobilisant différentes variables dans de multiples modèles visuels à des niveaux d'agrégations variés, ce datascape permet de multiplier les points de vue faisant avancer notre enquête pas à pas. Notre chemin interprétatif se déroule ainsi en onze étapes utilisant les divers outils d'analyse visuelle de la plateforme[2] .
Bien que nous ne prétendons pas ici apporter une contribution définitive à propos des conséquences de la guerre de Sept Ans sur l'économie française, le déroulé de notre cas d'usage complémente utilement la littérature sur le sujet .

Étape 1: vérifier la couverture temporelle du corpus



Figure 1 : nombre de flux de commerce par direction suivant les sources locales (vue Metadata) - http://toflit18.medialab.sciences-po.fr/datapaper/figure_9_step_1

La vue metadata a été conçue pour permettre aux chercheurs d'étudier finement la couverture du corpus. Nous étudions la période 1750-1770 afin d'observer le commerce avant, pendant et après la guerre de Sept Ans. La figure 1 montre le nombre de points de données disponibles dans les sources locales qui décrivent le commerce par direction[3]. Sept directions sont présentes dans le corpus sur la période à étudier : Marseille, Nantes, La Rochelle (sauf 1755 et 1763), Bordeaux, Rouen (très partiellement), Rennes et Bayonne.

Étape 2: vérifier la présence de flux de commerce avec les territoires du “Canada”

Pour isoler le commerce avec un partenaire commercial en particulier, il faut étudier la liste des noms des partenaires commerciaux mentionnés dans les sources. Les transcriptions contiennent 1050 noms de partenaires différents. Nous avons conçus un système de classification hiérarchique qui permet de réduire cette hétérogénéité de différente manière tout en explicitant les choix opérés.

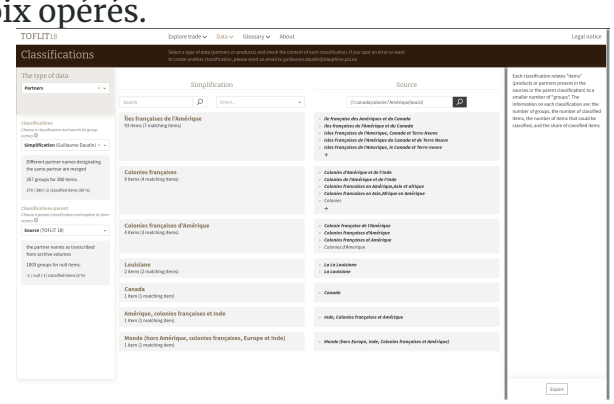


Figure 2 : partenaires commerciaux de la France contenant les termes canada ou colonie*amerique ou louisian (vue Classifications) - http://toflit18.medialab.sciences-po.fr/datapaper/figure_10_step_2

ans les transcriptions des sources il existe des références à "Canada", "Louisiane" et "Colonies d'Amérique" (Figure 2). Celles-ci sont agrégées en différents groupes dans la classification supérieure appelée "Simplification" : “Îles françaises de l’Amérique”, “Colonies Françaises”... Aucun de ces groupes ne correspond exactement à notre besoin d'isoler la Nouvelle France. Bien qu'il soit techniquement possible de créer une classification spécifique à un besoin de recherche, il serai dans notre cas méthodologiquement impossible d'isoler le Canada qui est presque systématiquement associé dans les sources aux “Îles Françoises de l’Amérique”.

Étape 3 : identifier le bon partenaire commercial “Îles françaises de l’Amérique”

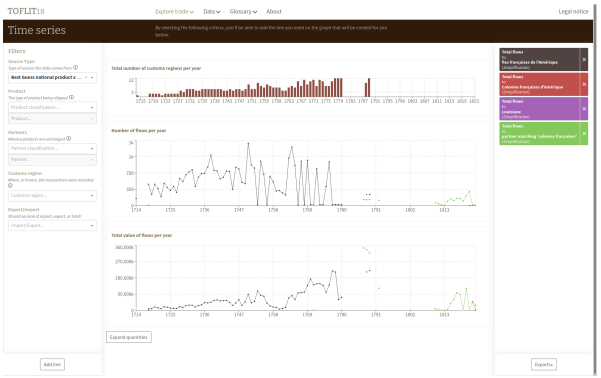


Figure 3 : comparaison en volume et valeur du commerce avec les partenaires "Îles françaises de l'Amérique", "Colonies françaises de l'Amérique", "Louisiane", "Colonies françaises". (vue Time series) - http://toflit18.medialab.sciences-po.fr/datapaper/figure_11_step_3

La Figure 3 montre que parmi les différents partenaires commerciaux identifiés précédemment, seules les “Îles françaises de l’Amérique” ont été mentionnées dans les sources pendant la guerre de Sept Ans.

Étape 4 : vérifier la position du partenaire dans la structure globale du commerce

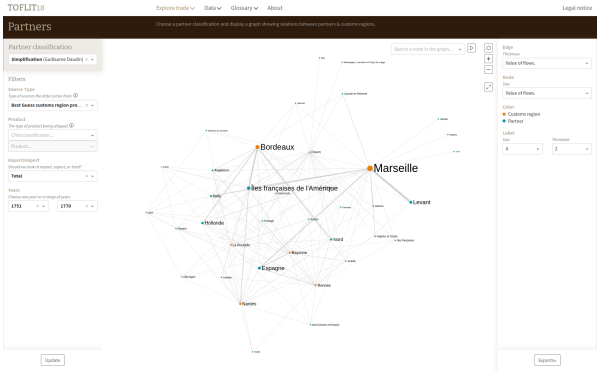


Figure 4 : réseaux des partenaires commerciaux de la France 1751-1770 (vue Locations) - http://toflit18.medialab.sciences-po.fr/datapaper/figure_12_step_4

Nous pouvons représenter l'importance relative de ce partenaire commercial dans le commerce de la France en visualisant le réseau des flux entre les directions et les partenaires commerciaux (Figure 4). En utilisant les valeurs des flux comme variable graphique du poids des liens et de la taille des nœuds, le partenaire “Îles françaises de l’Amérique” se démarque comme le principal partenaire de la France sur la période (en valeur). Les trois directions qui commercent le plus avec ce partenaire sont Marseille, Bordeaux et Nantes.
Les “Îles françaises de l’Amérique” semblent bien couvrir l'ensemble du commerce avec les colonies en Amérique. Mais le Canada n'est pas le seul territoire français concerné. Pour l'isoler nous allons devoir nous intéresser aux noms des marchandises.

Étape 5 : confirmer que les “Îles françaises de l’Amérique” comportent d'autres territoires que le Canada

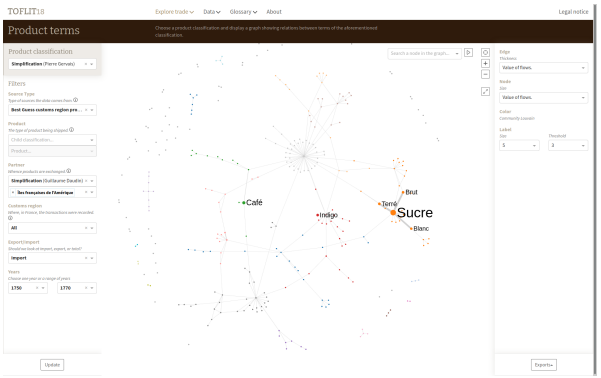


Figure 5 : réseaux des termes des produits importés par la France depuis les “Îles françaises de l’Amérique” 1750-1770 (vue Product terms) - http://toflit18.medialab.sciences-po.fr/datapaper/figure_13_step_5

La Figure 5 représente le réseau des termes utilisés dans les noms des produits importés depuis les “Îles françaises de l’Amérique” entre 1750 et 1770. Un lien entre deux termes indique qu'ils sont juxtaposés dans un nom de produit. Les noms de produits utilisés dans cette représentation sont issues de la classification "Simplification" qui agrège des noms de produits retranscrits des sources suivant le même système de classification hiérarchique que pour les partenaires commerciaux. En assemblant tous les liens de termes dans un même réseau on construit un espace sémantique qui se structure en fonction des mots utilisés dans différents contextes. Les nœuds de même couleur sont plus souvent associés ensemble qu'avec le reste du réseau des nœuds : ils forment un groupe sémantique[4] .
Seuls les labels des nœuds les plus importants en valeur sont mentionnés. Les quatre secteurs les plus importants en valeur sont le sucre (Sucre - Terré - Brut - Blanc), le café, les matières premières liées à l’industrie textile (Laine - Cotton - Indigo) et poisson (Morue - Séchée). Les trois premiers ne sont pas produits et très peu probablement réexportés par les territoires français du Canada mais plutôt par les possessions françaises dans les Caraïbes.

Étape 6 : construire une classification thématique des produits ciblant le “Canada”

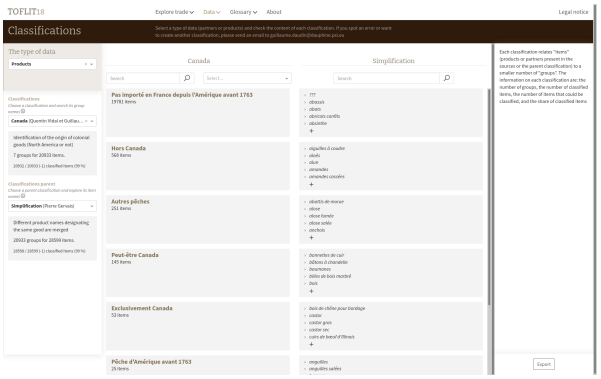


Figure 6 : La classification "Canada" (vue Classifications) - http://toflit18.medialab.sciences-po.fr/datapaper/figure_14_step_6

Pour isoler le commerce avec le Canada, nous avons donc construit une classification (nommée "Canada") qui classe les produits dans quatre groupes : "Hors Canada", "Exclusivement Canada", "Peut être Canada" et "intraitable"[5] . Cette classification est spécifique à une question de recherche. Elle se base comme de nombreuses autres classifications de même type dans notre système hiérarchique sur la classification "simplification". Cette dernière est dite générique car elle a été conçue pour couvrir de nombreux usages de recherche tout en permettant de réduire la trop grande variabilité des noms de produits en agrégeant notamment les cas de synonymies.

Étape 7 : isoler les échanges commerciaux avec le Canada

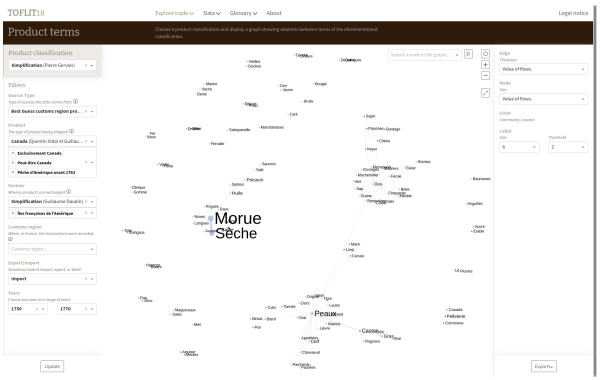


Figure 7 : réseau des termes des produits importés depuis le Canada, 1750-1751 (vue Product terms) - http://toflit18.medialab.sciences-po.fr/datapaper/figure_15_step_7

En filtrant les importations venant des “Îles françaises de l’Amérique” par les noms de marchandises en utilisant les catégories "Exclusivement Canada" et "Peut être Canada", il devient possible d'isoler les flux de commerce du Canada. Ce système est imparfait car il repose sur une liste de produits. C’est toutefois la seule possibilité d'agréger les flux de commerce d'une manière plus précise que ce que les noms des partenaires utilisés dans les sources permettent.
On observe sur le réseau de termes des importations avec le "Canada" (Figure 7) que les trois secteurs de marchandises les plus importants en valeur sont la morue, les peaux et le bois.

Étape 8 : analyser la répartition en France du commerce avec le “Canada”

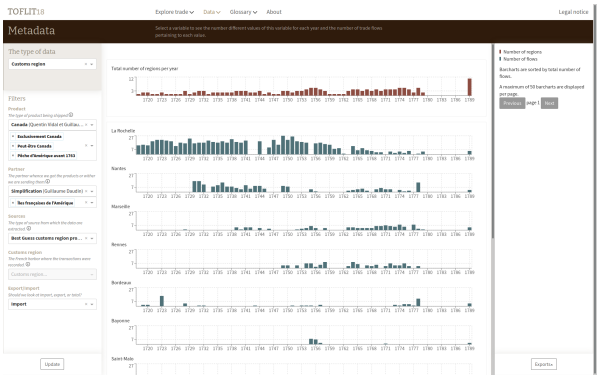


Figure 8 : Répartition par direction des imports depuis le Canada (vue Metadata) - http://toflit18.medialab.sciences-po.fr/datapaper/figure_16_step_8

Nous pouvons alors réutiliser cette méthode de filtrage du commerce avec le Canada dans la vue Metadata pour identifier les directions commerçant le plus (en nombre de flux[6]) avec ce dernier (Figure 8). Cette répartition nous permet de repérer les directions ayant le plus régulièrement importé depuis le Canada sur la période qui nous intéresse : La Rochelle, Nantes, Marseille et Rennes.

Étape 9 : comparer les plus grands ports importateurs de produits du Canada

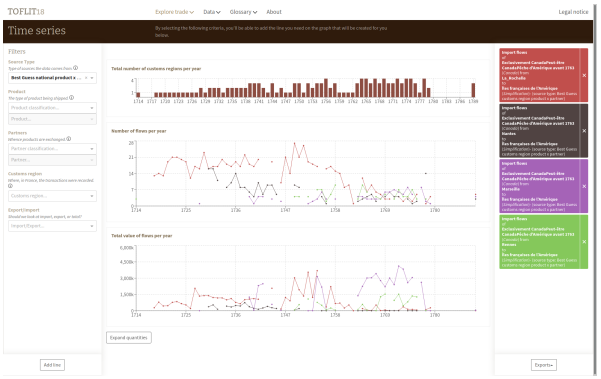


Figure 9 : évolutions des imports depuis le Canada de La Rochelle, Nantes, Marseille et Rennes - http://toflit18.medialab.sciences-po.fr/datapaper/figure_17_step_9

Nous examinons alors l'évolution du commerce de ces directions cette fois-ci en valeur. La Rochelle a été le plus important importateur de produits du Canada en valeur et en nombre de flux jusqu'en 1761 (Figure 9). Sur certaines années (1739, 1741, 1742 et 1754), les importations de Marseille sont plus importantes en valeur[7]. Les importations de La Rochelle déclinent à partir de 1750. Au contraire, Marseille dont les importations étaient sporadiques avant 1760 devient ensuite le plus grand importateur de produits "canadiens" en France. On remarque également que les importations de Rennes augmentèrent également après la guerre.

Étape 10 : révéler la spécialisation des ports

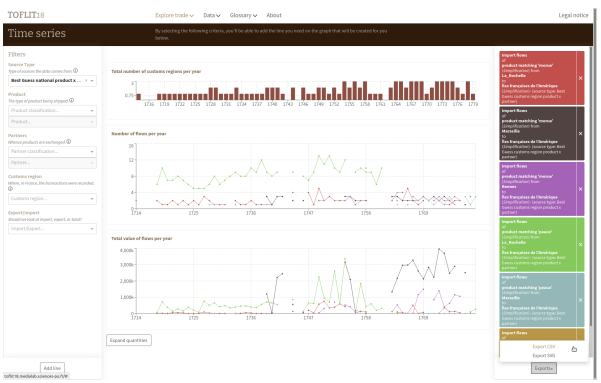


Figure 10 : évolution des imports de "morue" et de "peaux" depuis les "Îles Françaises de l'Amérique" par La Rochelle, Marseille et Rennes (vue Time series) - http://toflit18.medialab.sciences-po.fr/datapaper/figure_18_step_10

L'augmentation des valeurs d'importations depuis le Canada à Marseille et Rennes après la guerre de Sept Ans est principalement due aux importations de Morue (Figure 10). En effet le traité de Paris de 1763 permet à la France de conserver l'île de Saint-Pierre-et-Miquelon et un droit de pêche en Terre-Neuve. L'effondrement des importations de La Rochelle depuis le Canada correspond à la perte des territoires continentaux d'où proviennent une part importante des peaux (la perte des territoires allèrent de paire avec l'interdiction de commerce).

Étape 11 : confirmer un impact significatif sur le commerce de La Rochelle

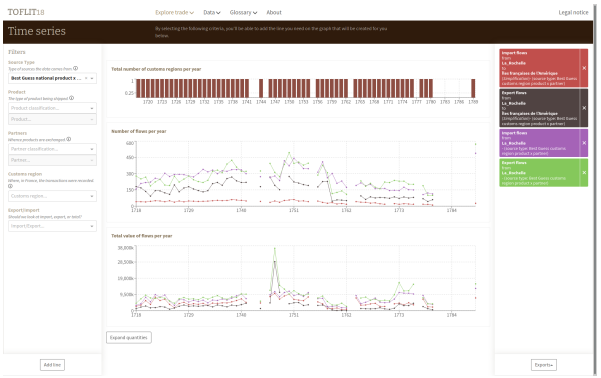


Figure 11 : évolution du commerce de La Rochelle total et avec les "Îles françaises de l'Amérique" (vue Time series) - http://toflit18.medialab.sciences-po.fr/datapaper/figure_19_step_11

Les visualisations permettent également de constuire des requêtes pour ensuite exporter des données. La Figure 11[8] montre qu'il est possible d'exporter en CSV les données de commerce total et avec les "Îles françaises de l'Amérique" de La Rochelle. À partir de ces données on peut calculer dans un autre logiciel diverses statistiques comme la part des “Îles françaises de l’Amérique” dans les importations et exportations de La Rochelle sur les cinq années disponibles avant 1756 et après 1763. La moyenne de la part du commerce des "Îles françaises de l'Amérique" dans les importations de La Rochelle est passé de 75% avant la guerre à 63% après (-12 points). Pour les exportations cette part passe de 47% à 20% (-27 points). Pour Bordeaux, Nantes et Marseille l’évolution de la part des "Îles françaises de l'Amérique" dans les exportations est respectivement de -8, +4 et +3 points .
Ainsi bien que la guerre de Sept Ans ne soit pas considérée comme ayant eu un impact important sur le commerce de la France , elle a bien eu un effet significatif sur La Rochelle et son économie locale qui perdit un marché important .

Conclusion

Ce récit d'exploration visuelle ne suffit pas à résoudre définitivement notre étude. Il devrait être complété par un travail complémentaire en croisant d'autres sources sur La Rochelle ou en calculant des statistiques confirmatoires. Il nous a en revanche permis d'illustrer les capacités de ce datascape via l'investigation d'une problématique en histoire économique et :

- de vérifier la couverture du corpus et les directions que l'exploration peut prendre par une documentation visuelle (vue Metadata);
- d'adapter l'hétérogénéité des sources à nos besoins par un système de classifications hiérarchiques (vue Classifications et filtrage);
- de bénéficier de moyens d'analyse visuelle révélant des effets temporels (vue Time series), structurels (vue Locations) et thématiques (vue Product terms);

- d'identifier des pistes de réponse à notre question de recherche et les données nous permettant de poursuivre l'enquête par des outils d'analyse plutôt que d'exploration.

En détaillant les étapes de notre enquête à propos de la perte du Canada, nous avons mis en avant des fonctionnalités qui permettent de relever les défis de l'analyse quantitative historique : prise en compte des effets de construction des données (disponibilité et hétérogénéité), adaptabilité des méthodes d'agrégation et des moyens d'exploration aux questions de recherche, importance de la circulation entre différentes vues (longitudinale, structurelle…) pour isoler les effets statistiques pertinents. Ces dynamiques exploratoires favorisant la multiplication des points de vue nous semblent particulièrement adaptées aux recherches en humanités qui mobilisent des sources souvent complexes et qui exigent de concilier quantification et complexité interprétative .
Construire un tel instrument demande un investissement important et une très forte collaboration entre sciences sociales, ingénierie et design . L'intérêt de ne se résume d'ailleurs pas à la mise à disposition d'un outil et d'un jeu de données[9]. Nous avons été contraints de penser transcription, modélisation, exploitation quantitative et visuelle d'un seul tenant. Le processus de conception et production du datascape a largement contribué à améliorer notre maîtrise de la source. Surtout, cette collaboration a mis en cause nos habitudes et fait circuler savoir-faire et perspectives.

1. Cet article résulte d'une communication prévue pour le colloque Humanistica 2020 annulé suite à la pandémie COVID19. Il a été écrit pour être lu sur le web grâce au logiciel d'écriture savante Fonio

2. Nous précisons que toutes les figures sont des copies d'écran du datascape. Seules les quelques statistiques de la dernière étape ont été calculées en dehors.

3. On note dans la figure 1 que l'année 1750 compte beaucoup plus de directions. C'est en effet l'année que nous avons choisi pour permettre une couverture complète du territoire français.

4. Plus précisément la couleur résulte de l'
algorithme de détection de communauté par modularité selon Louvain.

5. La première version de cette classification a été produite par Quentin Vidal pour son mémoire de Master.

6. Le nombre de flux reflète indirectement la diversité (et non la valeur) des marchandises échangés car les flux de commerce sont enregistrés par marchandise. Pour un même partenaire commercial, le plus grand nombre de marchandises différentes, le plus grand nombre de flux.

7. À propos de l'activité du port de Marseille au XVIIIe siècle voir (Carrière, 1973)

8. Cette figure suggère aussi d'aller regarder de plus prêt l'année 1747 du commerce de La Rochelle qui connaît un pic en valeur suspect. Peut être une erreur de transcription à débusquer.

9. L'outil présenté dans ce papier est librement accessible
http://toflit18.medialab.sciences-po.fr, son code source est libre et ouvert. Les données (en cours de documentation sous la forme d'un data package) est disponible sur ce dépôt en attendant sa publication.

References

Baugh, D. A. (2011). The global Seven Years War, 1754-1763: Britain and France in a great power contest. Harlow, England; New York: Longman.
Carrière, C. (1973). Négociants marseillais au XVIIIe siècle : contribution à l'étude des économies maritimes. Marseille.
Charles, L., & Daudin, G. (2011). La collecte du chiffre au XVIIIe siècle : le Bureau de la balance du commerce et la production des données sur le commerce extérieur de la France. *Revue d'histoire moderne & contemporaine*, *n° 58-1*(1), 128-155. Consulté à l'adresse https://www.cairn.info/revue-d-histoire-moderne-et-contemporaine-2011-1-page-128.htm
Clark, J. G. (1981). La Rochelle and the Atlantic economy during the eighteenth century. Baltimore: Johns Hopkins University Press.

Drucker, J. (2011). Humanities Approaches to Graphical Display. *Digital Humanities Quarterly*, *5*(1). Consulté à l'adresse http://www.digitalhumanities.org/dhq/vol/5/1/000091/000091.html
Girard, P. (2018). Collaborative Humanities. Maison européenne des sciences de l'homme et de la société de Lille, FR. Consulté à l'adresse http://spire.sciencespo.fr/hdl:/2441/7b3nh3cnt99qjb9b060q4cbumt
Girard, P., Jacomy, M., & Leclercq, C. (2017). *Explorer associations et transformations à travers des datascapes*. Paris, France. Consulté à l'adresse http://spire.sciencespo.fr/hdl:/2441/2solvroumm9kgpsu49vb8pphju
Girard, P., & Plique, G. (2019). *Using networks to study 18th century French trade*. Bruxelles, Belgium. Consulté à l'adresse https://fosdem.org/2019/schedule/event/graph_french_trade_study/
Latour, B., Jensen, P., Venturini, T., Grauwin, S., & Boullier, D. (2012). 'The whole is always smaller than its parts' – a digital test of Gabriel Tarde's' monads. *The British Journal of Sociology*, *63*(4), 590-615. Consulté à l'adresse http://onlinelibrary.wiley.com/doi/10.1111/j.1468-4446.2012.01428.x/abstract
Martinetti, B. (2013). Les négociants de La Rochelle au xviiie siècle. Histoire. Rennes: Presses universitaires de Rennes. Consulté à l'adresse http://books.openedition.org/pur/116217
Mathieu Bastian, Sebastien Heymann, & Mathieu Jacomy. (2009). Gephi: An Open Source Software for Exploring and Manipulating Networks. *International AAAI Conference on Web and Social Media; Third International AAAI Conference on Weblogs and Social Media*. Consulté à l'adresse https://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154
Riley, J. C. (1986). *The Seven Years War and the Old Regime in France: the economic and financial toll*. Princeton, N.J.: Princeton University Press.
Tukey, J. W. (1997). *Exploratory data analysis*. Addison-Wesley series in behavioral science Quantitative methods. Reading, Mass.: Addison-Wesley.