

Le projet RICardo : une nouvelle source sur l'histoire de la mondialisation commerciale

Béatrice Dedinger

L'auteur

Béatrice Dedinger est chercheur au Centre d'histoire de Sciences Po. Ses recherches ont d'abord porté sur l'histoire commerciale de l'Allemagne. Elle a notamment publié : « Exploring Trade Globalization in the Long Run: The RICardo Project » (*Historical Methods*, 2017) ; « Trade statistics of the Zollverein » (*Revue de l'OFCE*, 2015) ; « The Franco-German trade puzzle : an analysis of the economic consequences of the Franco-Prussian war » (*Economic History Review*, 2012) ; « L'avenir commercial de l'Allemagne » (*Revue d'Allemagne*, 2013). Béatrice Dedinger est coordinatrice du projet RICardo depuis son lancement et est reconnue comme une spécialiste des statistiques commerciales des XIX^e et XX^e siècles.

Résumé

Le projet RICardo, démarré en 2004, a conduit à la création d'un corpus de données et d'un site internet ouvert au public en décembre 2017. Il est consacré aux relations commerciales bilatérales des pays du monde sur la période 1800/30-1938. Cet article a pour but de présenter les grandes étapes du projet, de la collecte des données à l'élaboration d'un outil numérique permettant de les traiter ; d'éclairer le lecteur sur les caractéristiques des statistiques commerciales historiques ; et de montrer en quoi RICardo offre de nouveaux outils pour la recherche et l'enseignement sur l'histoire du commerce et de la mondialisation commerciale.

Mots clés : analyse exploratoire des données ; bases de données commerciales ; histoire du commerce international ; mondialisation commerciale ; statistiques commerciales historiques.

Abstract

"The RICardo project: a new source on the history of trade globalization"

Launched in 2004, the RICardo project has resulted in the creation of a corpus of data and an internet site that was opened to the public in December 2017. It is devoted to the bilateral trade relations of the world's countries for the period 1800/30-1938. The present article has several aims: to present the major stages of this project, from the collection of data to the development of a digital tool for handling it; to inform the reader as to the characteristics of historical trade statistics; and to show in what ways RICardo offers new tools for research and teaching in the history and globalization of trade.

Key Words: Exploratory Data Analysis; Trade Databases; History of International Trade; Trade Globalization; Historic Trade Statistics.

Pour citer cet article : « Le projet RICardo : une nouvelle source sur l'histoire de la mondialisation commerciale », *Histoire@Politique*, n° 34, janvier-avril 2018 [en ligne, www.histoire-politique.fr]

L'idée du projet RICardo (RICardo pour Research on International Commerce) est née fin 2004¹. Partant du constat que les bases de données commerciales élaborées jusque-là restaient limitées aux pays occidentaux et développés et ne remontaient pas au-delà des années 1870, un groupe de chercheurs de Sciences Po² a envisagé de construire une base de données couvrant tous les pays du monde et la période la plus longue possible. Deux faits majeurs ont permis de passer de l'idée à la réalisation du projet. Tout d'abord, la découverte d'une série d'archives du commerce extérieur dans les fonds de la Bibliothèque nationale de France qui rappelait que des statistiques commerciales étaient enregistrées par certains États au moins depuis les années 1830. Les *Annales du commerce extérieur*, publiées à partir de 1829, rassemblent les rapports rédigés par les consuls de France établis dans les différents pays du monde et contiennent des informations statistiques d'une grande richesse. Le second élément déclencheur a été l'obtention d'un soutien financier de l'Agence nationale de la recherche suffisamment important pour pouvoir se lancer dans l'entreprise³. Après plus de dix ans de travail, le projet RICardo peut enfin se prévaloir, depuis décembre 2017, d'une réalisation concrète avec le lancement officiel du site internet « *The RICardo Project* »⁴. Il est heureux que la date de lancement coïncide avec le bicentenaire de la publication du célèbre ouvrage de David Ricardo, *On the Principles of Political Economy and Taxation*.

Créé à l'origine par des historiens économistes pour des besoins de recherche en histoire économique, le projet a peu à peu mobilisé des chercheurs de différentes disciplines incluant, outre l'histoire économique, la géographie, les sciences politiques, la statistique et les humanités numériques. Le caractère inter- et transdisciplinaire du projet est à souligner. Il se retrouve dans la composition du comité RICardo⁵ et il a pour effet d'élargir considérablement l'éventail des utilisations possibles ainsi que le public visé.

Cet article a pour but de faire connaître le projet RICardo aux lecteurs de la revue *Histoire@Politique*. L'univers des statistiques commerciales est encore réservé à un cercle restreint de spécialistes. Ces données peuvent cependant intéresser un plus large public par les questions qu'elles posent, qu'elles soient de nature historique, économique, voire philosophique, et par l'usage ludique que permettent aujourd'hui les technologies numériques. Un premier point présentera le projet RICardo, son originalité, ce que la base de données possède de plus que les bases déjà existantes, les méthodes selon lesquelles elle a été construite, ce que propose le site internet. Un deuxième point portera sur les caractéristiques des statistiques commerciales historiques dont les défauts sont bien connus des spécialistes, mais qui restent un des meilleurs outils pour étudier l'histoire de la mondialisation et de la croissance. Le troisième point sera consacré aux perspectives de renouvellement de la recherche et de l'enseignement sur l'histoire de la mondialisation offertes par le projet RICardo.

¹ L'auteur tient à remercier Paul Girard, directeur technique du médialab de Sciences Po, pour ses commentaires. Paul Girard supervise toute la partie numérique du projet RICardo. Cet article a également bénéficié de la relecture critique de deux rapporteurs anonymes.

² De la Chaire Finances internationales alors dirigée par Marc Flandreau.

³ Projet ANR (2007-2010), « Comprendre la mondialisation commerciale ». Depuis 2011, le projet RICardo a pu être financé grâce à la procédure d'appel à projets instaurée par Sciences Po.

⁴ <http://ricardo.medialab.sciences-po.fr/>. Un carnet blog a également été ouvert : <https://ricardo.hypotheses.org>.

⁵ Présentation du comité et de ses premiers travaux à l'adresse suivante : https://f-origin.hypotheses.org/wp-content/blogs.dir/4050/files/2017/11/RIC_Workshop_2017_Report.pdf

Brève présentation du projet RICardo⁶

RICardo comprend une base de données et un site internet. Dans le projet retenu en 2007 par l'ANR, le principal objectif était la création d'une base de données. Le déploiement soudain des technologies numériques à la fin des années 2000 a conduit, en 2013, à un renouvellement des objectifs ainsi que de l'équipe pour permettre l'élaboration d'un site internet. RICardo rassemble des statistiques commerciales de tous les pays du monde sur une période de plus de cent ans, allant de la fin du XVIII^e siècle jusqu'en 1938⁷. En raison des lacunes dans les sources sur les périodes de guerre et d'hyperinflation, il a été décidé au départ de ne pas couvrir les périodes 1914-1923 et 1939-1945. Ce choix a été rediscuté par la suite et, pour commencer, une nouvelle collecte de données a été entreprise sur la période 1914-1923. Dans sa deuxième version (novembre 2017), la base comprend 300 000 données. Une donnée correspond à un flux d'exportation ou d'importation, total ou bilatéral. Les exportations/importations totales correspondent à la somme des marchandises exportées/importées par un pays sur une période ; les exportations/importations bilatérales désignent la répartition géographique du commerce total d'un pays sur une période. Le site internet propose des visualisations graphiques de ces données, allant d'une vision agrégée du commerce des pays (évolution du commerce mondial) à une vision désagrégée (évolution des échanges entre deux pays). Il est à noter que la base ne contient pas de statistiques sur les échanges de produits.

Petit historique des bases de données commerciales

Les statistiques commerciales sont parmi les premières à avoir été collectées par les États, pour des raisons avant tout fiscales. En effet, le prélèvement d'un droit de douane sur une marchandise importée ou exportée nécessite de connaître son poids et/ou sa valeur. Les premiers registres de douane remonteraient au XVI^e siècle⁸. Ce sont des manuscrits qui fournissent des relevés très détaillés des échanges de produits et nécessitent un important travail de transcription, de classification et d'agrégation avant de pouvoir être utilisés par les chercheurs. La publication régulière de tableaux du commerce sous forme d'imprimés a commencé au début du XIX^e siècle dans plusieurs pays d'Europe, d'Amérique et d'Asie⁹. On y trouve des tableaux résumés et agrégés des échanges par pays et par produits qui permettent de

⁶ Pour plus de détails sur les questions méthodologiques et la dimension numérique du projet, voir Dedinger, Béatrice et Paul Girard, « Exploring trade globalization in the long run: The RICardo project », *Historical Methods : A Journal of Quantitative and Interdisciplinary History*, 2017, 50(1), p. 30-48.

⁷ La base contient relativement peu de données avant les années 1830. Plusieurs projets sont actuellement menés qui visent à reconstituer des séries de statistiques commerciales sur le XVIII^e siècle. Pour une revue détaillée de ces projets, voir Loïc Charles et Guillaume Daudin (dir.), *Eighteenth-Century International Trade Statistics. Sources and Methods*, OFCE, Revue de l'OFCE, 2015. Noter qu'après la Seconde Guerre mondiale, les organismes internationaux, FMI, GATT et Nations Unies, seront chargés de centraliser et de publier les statistiques commerciales des pays. Des bases commerciales historiques, couvrant toute la période d'après-guerre, sont aujourd'hui disponibles et en libre accès (<http://data.imf.org>; <http://comtrade.un.org>, lien consulté le 15/02/2018).

⁸ Cf. L. Charles et G. Daudin (dir.), *Eighteenth-Century International Trade Statistics...*, *op. cit.*, p. 24-25. Les plus anciens registres aujourd'hui inventoriés sont ceux des États italiens (République de Gênes, Royaume de Naples, États pontificaux) et du Sund (détroit danois au passage duquel les bateaux étrangers devaient payer un droit de douane).

⁹ Voir la liste des sources données sur le site RICardo. Cette liste est susceptible d'être complétée.

collecter directement les données sans avoir à passer par un travail préalable de reconstitution. Des compilations de ces statistiques nationales sont publiées par le Royaume-Uni à partir des années 1850 dans deux séries de *Statistical Abstract*, et par la France dans la série *Annales du commerce extérieur*. Après 1918, la centralisation des statistiques commerciales des pays du monde sera confiée à la Société des Nations¹⁰.

Toutes ces sources commenceront à être exhumées par les historiens économistes dans les années 1950. Les travaux les plus connus sont ceux de Lewis, Maddison et Mitchell qui ont pour but d'étudier l'évolution du commerce mondial¹¹. Ces chercheurs ont reconstitué des séries historiques d'exportations et d'importations totales des pays qu'ils ont publiées sur papier. Le progrès technique et la création de logiciels de traitement de données ouvrent la voie à une génération de bases de données plus exhaustives à partir des années 1990. Simultanément, de nombreux chercheurs se sont intéressés aux causes et aux conséquences de la mondialisation qu'ils ont analysées à l'aide du modèle de gravité¹². Son utilisation nécessitant des bases de données de taille relativement importante, chaque chercheur s'est mis à construire sa propre base de données bilatérales pour les besoins d'un travail personnel. Le premier chercheur à avoir entrepris de créer une base historique ayant une large couverture spatiale et temporelle, non seulement pour ses propres besoins mais aussi pour ceux de la communauté scientifique, est Katherine Barbieri. Dans le but d'étudier les liens entre commerce et conflits, cette politiste américaine a construit et mis en ligne une base de données bilatérales sur la période 1870-à nos jours¹³. D'autres bases ont été élaborées par la suite : centrée sur l'entre-deux-guerres¹⁴, faisant la synthèse de bases existantes¹⁵, proposant de nouvelles séries d'exportation et d'importations mondiales sur 1800-1938¹⁶.

Quelle est la valeur ajoutée du projet RICardo par rapport à l'ensemble de ces travaux ? Tout d'abord, RICardo est un projet évolutif. La base de données n'est pas figée, il est prévu de continuer à l'enrichir grâce à la collecte de nouvelles sources et données. C'est aussi un projet collaboratif. L'esprit de RICardo est d'œuvrer à la fédération des compétences en matière de statistiques commerciales historiques et de créer un pôle de référence dans ce domaine. C'est pourquoi il a été proposé aux

¹⁰ À partir de 1948, le FMI centralise et publie les données de commerce bilatéral. Le GATT/OMC se concentre sur les statistiques d'échanges de produits.

¹¹ Arthur Lewis, « The Rate of Growth of World Trade 1830-1973 », dans Sven Grassman et Erik Lundberg (dir.), *The World Economic Order. Past and Prospects*, Macmillan, London, 1981, p. 29-30 ; Angus Maddison, « Growth and Fluctuation in the World Economy, 1870-1960 », *Banca Nazionale del Lavoro Quarterly Review*, 1962, 61, p. 127-195 ; Brian R. Mitchell, *International Historical Statistics*, 3 vol., 2007 (dernière édition).

¹² Ce modèle explicatif de la structure des échanges bilatéraux stipule que les échanges entre deux pays sont déterminés par leur poids économique (le PIB) et la distance entre eux ; d'autres variables (par exemple, l'existence d'un accord commercial ou monétaire) permettent d'améliorer le pouvoir explicatif du modèle.

¹³ La quatrième version de cette base a été publiée en 2016. Barbieri, Katherine et Omar M. G. Omar Keshk. 2016. *Correlates of War Project Trade Data Set Codebook, Version 4.0*. Online: <http://correlatesofwar.org> [lien consulté le 15/02/2018].

¹⁴ Joanne Gowa et Raymond Hicks, « Commerce and Conflict: New Data about the Great War », *British Journal of Political Science*, 2015, DOI: <http://dx.doi.org/10.1017/S0007123415000289> [lien consulté le 15/02/2018].

¹⁵ Michel Fouquin et Jules Hugot, « Two Centuries of Bilateral Trade and Gravity Data: 1827-2014 », *Working Paper CEPII* 2016-14.

¹⁶ Giovanni Federico et Antonio Tena-Junguito, « World Trade, 1800-1938: a New Data-set », *EHES Working Papers in Economic History*, 2016, 93.

historiens économistes, Giovanni Federico et Antonio Tena, d'intégrer au site RICardo les nouvelles séries de commerce mondial qu'ils ont récemment estimées. Cette dimension collaborative est entretenue au sein du comité RICardo, par des liens étroits avec d'autres projets dédiés aux statistiques commerciales historiques, et grâce aux retours des chercheurs qui utilisent la base de données. RICardo est enfin un projet numérique, ce qui reste sa dimension la plus innovante. Le passage de la collecte de données à la création du site internet s'est fait en plusieurs étapes qui seront brièvement décrites.

De la collecte des données au site internet

La collecte des données est l'étape la plus exigeante en main-d'œuvre et en budget. Depuis le début du projet, une douzaine de personnes (étudiants, chercheurs) ont travaillé à la saisie des statistiques¹⁷. Au départ, la collecte s'est faite directement en bibliothèque ou à partir de photocopies. Aujourd'hui, la numérisation des documents est devenue systématique. Riitta Hjerppe de l'université d'Helsinki et Lars-Fredrick Andersson de l'université d'Umeå ont apporté leur contribution pour les statistiques finlandaises et suédoises. Au départ du projet, il a été décidé de privilégier les sources secondaires (compilations de statistiques commerciales nationales) et de ne pas renseigner la période 1914-1923 (souvent lacunaire dans les sources secondaires). Cette stratégie a été révisée en 2012, en même temps que la dimension et l'objectif du projet. Un travail de collecte de sources primaires (publications nationales de statistiques commerciales) et des données de la période 1914-1923 a été entrepris, toujours en cours à ce jour.

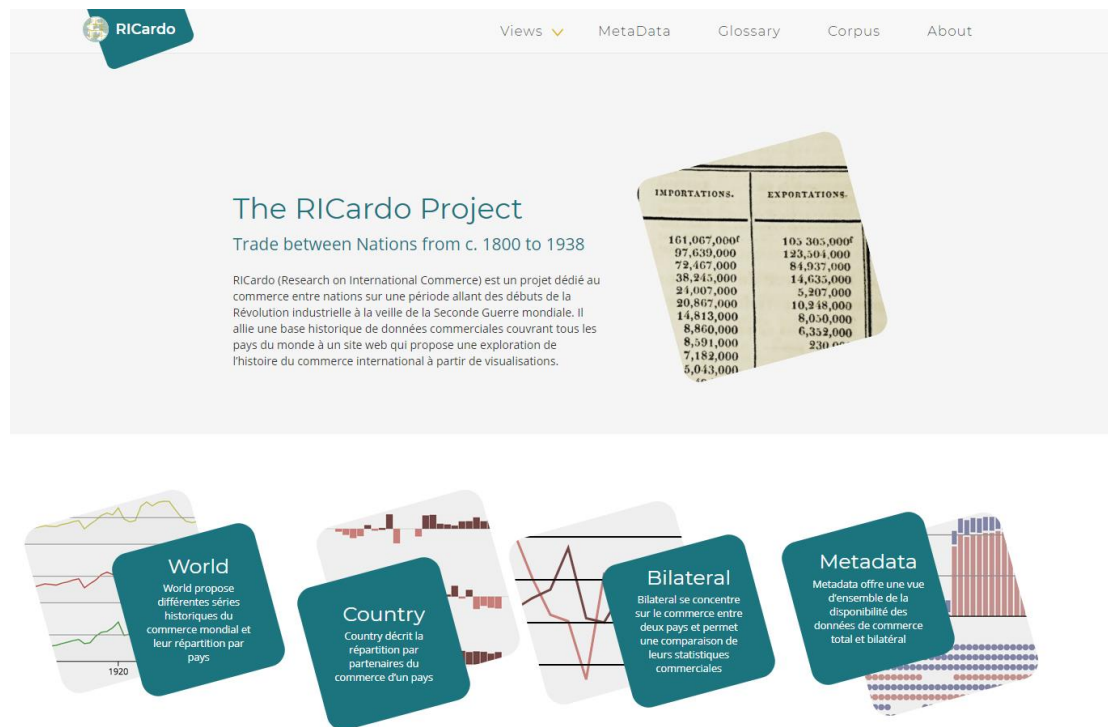
Une fois que les données ont été récoltées, l'étape suivante consiste à les rendre utilisables. Cette étape a nécessité les services d'un gestionnaire de base de données afin d'élaborer une structure de base de données relationnelle reliant les données originales à des tables de conversion¹⁸. Les noms des entités¹⁹, que l'on trouve dans la base d'origine en plusieurs langues et/ou mal orthographiés, doivent être nettoyés et homogénéisés. Les valeurs des flux commerciaux, qui sont exprimées en une centaine de monnaies différentes, ont été converties en une unité commune, la livre sterling²⁰. Il a fallu pour cela créer une base de taux de change qui comprend actuellement un peu plus de 7 000 données sur la période 1800-1938. C'est ainsi que le projet RICardo a produit un corpus de données qui inclut une base historique de données commerciales, une base historique de taux de change, une table des entités et une table des sources. Ce corpus est mis à jour chaque fois que de nouvelles données sont intégrées.

¹⁷ La première version de la base de données a été créée et gérée sous Microsoft Access. La décision de créer un site internet a conduit à revoir ce choix et à convertir la base de données au format SQLite.

¹⁸ Pour plus de détails sur cette étape, voir Béatrice Dedinger et Paul Girard, *op. cit.*, p. 46-48.

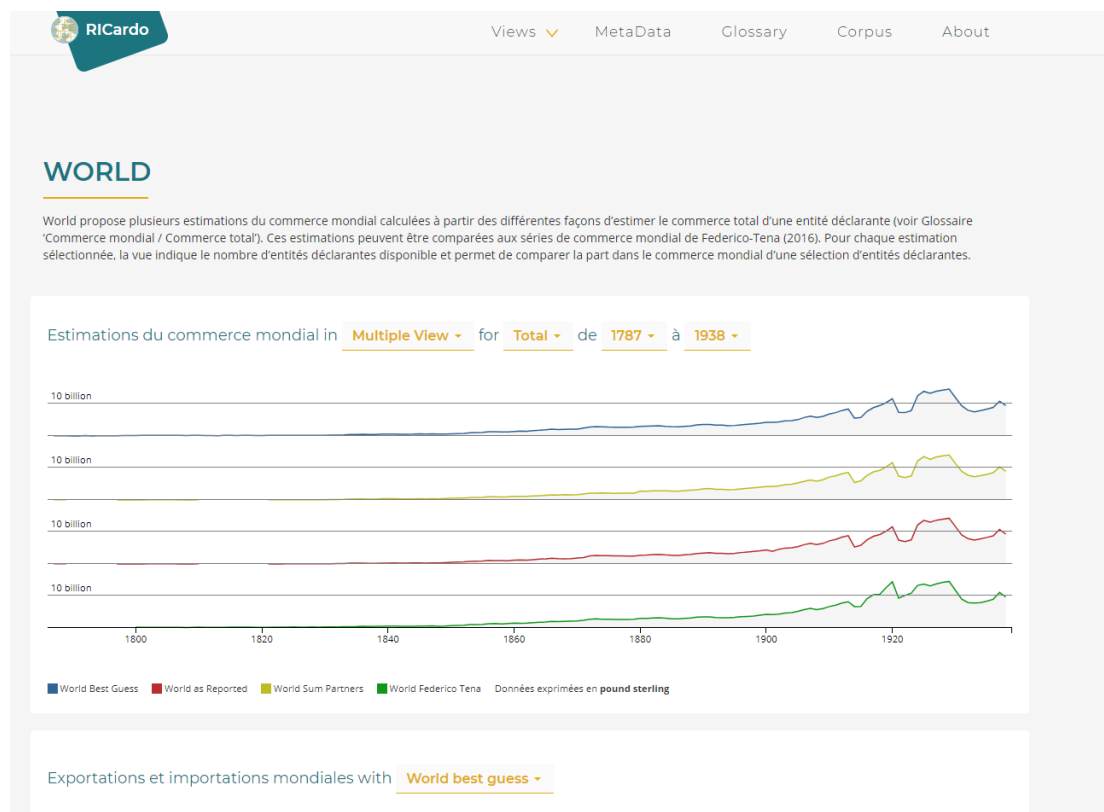
¹⁹ Le terme « entité » est utilisé plutôt que celui de « pays » car il n'est pas rare, dans les statistiques commerciales historiques, que l'entité commerciale ne corresponde pas à un État internationalement reconnu, mais se rapporte à un lieu, une région ou un groupe d'États.

²⁰ Tous les flux sont exprimés en valeurs courantes. Sur cette question, voir le rapport du comité RICardo : https://f-origin.hypotheses.org/wp-content/blogs.dir/4050/files/2017/11/RIC_Workshop_2017_Report.pdf [lien consulté le 15/02/2018].



Vue de la page d'accueil

L'ouverture d'un site internet n'avait pas été envisagée au départ. Cette étape a demandé cinq années supplémentaires de travail ainsi que de nouveaux investissements humains et financiers. Elle a profondément modifié l'objectif du projet. À l'origine, seul était visé un public de chercheurs spécialisés désireux de disposer facilement d'une masse de données importante pour nourrir des tests économétriques. La création d'un site nous a conduits à réfléchir à d'autres utilisations de la base par un public plus étendu. C'est la visualisation instantanée des données de son choix par l'utilisateur qui constitue le principal attrait d'un site. Trois visualisations sont proposées sur le site RICardo qui correspondent à différents niveaux d'agrégation des flux de commerce. Dans la vue *'World'*, l'utilisateur peut voir l'évolution du commerce du monde et des pays du monde du début du XIX^e siècle jusqu'en 1938. Plusieurs estimations du commerce total des pays sont calculées et comparées dont l'intérêt sera expliqué dans le dernier point. La vue *'Country'* se concentre sur le commerce d'un pays avec l'ensemble de ses partenaires. Elle permet de suivre l'évolution du commerce par pays et le poids des partenaires commerciaux. La dernière vue, *'Bilateral'*, montre le commerce entre deux pays et offre une comparaison des « flux miroirs ». Cette visualisation est la plus originale du site. Elle exploite le fait que la base RICardo fournit quatre flux bilatéraux pour chaque paire de pays. Un flux commercial est enregistré deux fois, par chacun des pays commerçants, dans deux sources différentes. On peut ainsi comparer les exportations de A vers B, telles qu'elles sont déclarées dans la source A, aux importations de B en provenance de A enregistrées dans la source B. C'est ce qu'il est convenu d'appeler un flux miroir. Théoriquement, les deux flux devraient être égaux, mais ce n'est jamais le cas dans la pratique. Cette question de la divergence des flux bilatéraux sera développée dans le point suivant.



Vue de la page « World »

L'intérêt premier du projet RICardo est de mettre à la libre disposition des chercheurs une base historique de données commerciales et d'ouvrir à un public de non-spécialistes le champ de l'utilisation des statistiques commerciales historiques. L'usage de ces statistiques nécessite cependant d'être averti sur la qualité des données qui est loin d'être parfaite.

Imperfection des statistiques commerciales

Le problème de la fiabilité des statistiques commerciales est posé au moins depuis le milieu du XIX^e siècle²¹. Un siècle plus tard, on peut encore lire ce constat sans appel d'un économiste renommé pour ses travaux sur l'erreur en statistiques économiques : « ceux qui écrivent sur le commerce extérieur, quelle que soit la période, devront assumer la charge de la preuve que les données sur les mouvements de marchandises sont suffisamment bonnes pour garantir la manipulation et le raisonnement dont elles sont habituellement l'objet²² ». Comment les historiens économistes experts en statistiques commerciales doivent-ils traiter cette question ? La réponse qui paraît la

²¹ Voir l'article du *Journal of the Statistical Society of London* publié en 1861, « Memorandum indicating some of the causes to which the discordancy of the commercial statistics of various countries is chiefly to be attributed ».

²² « *Writers on all phases of foreign trade will have to assume the burden of proof that the figures on commodity movements are good enough to warrant the manipulation and the reasoning to which they are customarily subject.* » Oskar Morgenstern, *On the Accuracy of Economic Observations*, Princeton, Princeton University Press, 1963, p. 180.

plus satisfaisante est de nature philosophique. C'est celle donnée par Henri Guitton dans la préface de l'édition française de l'ouvrage d'Oskar Morgenstern :

« Dès la première édition de 1950, j'avais été frappé par l'immense portée des réflexions que suscitait cet ouvrage. J'enseignais alors la Statistique. Le problème des erreurs me paraissait fondamental à faire comprendre, non pas pour dénigrer la Statistique, comme cela est arrivé si souvent d'une manière naïve et même un peu enfantine : ne pas utiliser des données statistiques sous prétexte qu'elles ne sont pas exactes ! C'est pourquoi je me suis employé dans ma petite sphère à habituer les esprits à travailler dans l'imperfection des saisies chiffrées des phénomènes, pour en tirer cependant le maximum de vérité. La philosophie des erreurs commande l'intelligence même de la Statistique²³. »

Tel est l'esprit du projet RICardo. Le but n'est pas de tendre vers une base « parfaite » en lissant les données, en les extrapolant ou en les complétant à l'aide des flux miroirs (cf. ci-dessous), ce qui eût été source de nouvelles erreurs. Une documentation a été mise sur le site afin d'informer l'utilisateur sur cette question de fiabilité et la vue '*Bilateral*' a été créée afin de l'aider à mieux comprendre l'imperfection des statistiques commerciales.

Comment sont élaborées les statistiques commerciales ?

Les statistiques commerciales sont des statistiques publiques établies par un État à partir des déclarations en douane sur la quantité et/ou la valeur des marchandises qui entrent sur et sortent du territoire douanier (qui ne coïncide pas toujours avec le territoire politique). C'est dans les ports que sont enregistrés ces échanges internationaux de marchandises²⁴. Toutes les déclarations sont recueillies par un office central de statistique (c'est une création relativement récente, à l'origine existaient des Bureaux de commerce ou des douanes) dont la fonction est de produire un état du commerce du pays avec l'extérieur. Pour cela, un travail doit être fait sur la valorisation et la classification des flux commerciaux. En effet, les premiers registres de douane ne donnent pas toujours d'information sur la valeur ou le prix des marchandises échangées²⁵. On en trouve des exemples dans nombre de registres établis à partir du XVI^e siècle²⁶. Le cas le plus remarquable est celui du Zollverein allemand qui a publié, de 1844 à 1871, un état du commerce dans lequel ne figurent que des données en quantité²⁷, ce qui ne permet pas de connaître la valeur des échanges totaux ou bilatéraux²⁸. La valorisation des flux commerciaux peut se faire suivant différentes méthodes. Les valeurs peuvent être estimées par les bureaux de

²³ Oskar Morgenstern, *L'illusion statistique. Précision et incertitude des données économiques*, Paris, Dunod, 1972, préface.

²⁴ Voir le projet Toflit18 qui vise à reconstituer le commerce extérieur de la France au XVIII^e siècle à partir des registres de commerce de tous les ports français (<http://toflit18.hypotheses.org>).

²⁵ Il faut faire une distinction entre prix et valeur. La valeur d'un flux de commerce est égale à la quantité de ce flux multiplié par le prix unitaire du bien. Pour reconstituer des séries de commerce en valeur lorsqu'ils ne disposent que des quantités, les chercheurs utilisent des séries de prix unitaires.

²⁶ Voir dans Loïc Charles et Guillaume Daudin, *op. cit.*, p. 24-25, le tableau récapitulatif par pays des informations que l'on trouve dans les premiers registres du commerce.

²⁷ Principalement en *Centner*. Jusqu'en 1840, 1 *Prussian Centner* = 100 kg ; par la suite, 1 *Zoll Centner* = 50 kg.

²⁸ Cf. Béatrice Dedinger, « Le problème des statistiques allemandes du commerce extérieur avant 1871 », *Histoire@Politique. Politique, culture, société*, n° 17, mai-août 2012 [en ligne : <https://histoire-politique.fr/index.php?numero=17&rub=autres-articles&item=67>], ; Béatrice Dedinger, « Trade Statistics of the Zollverein, 1834-1871 », dans Loïc Charles et Guillaume Daudin, *op. cit.*, 2015, p. 67-85.

commerce (en utilisant les prix d'une année donnée, en chargeant une commission d'experts d'évaluer les prix, en récoltant les prix auprès des chambres de commerce) ou elles peuvent être déclarées par les exportateurs et importateurs. Le système des valeurs estimées, considéré comme le moins fiable, a prévalu jusqu'à l'entre-deux-guerres. La classification des marchandises est, elle aussi, essentielle pour permettre l'utilisation des statistiques commerciales, que ce soit à des fins politique, économique ou scientifique. Des classifications très simples ont pu tout d'abord être proposées²⁹. Une nouvelle étape de l'histoire des classifications de marchandises est franchie en 1913 avec la publication d'une nomenclature internationale unifiée, la classification de Bruxelles³⁰. Depuis 1919, l'harmonisation et l'adaptation des nomenclatures de marchandises ont été une des missions confiées à la Société des Nations puis aux Nations Unies³¹.

Une autre dimension de l'élaboration des statistiques commerciales est la répartition des échanges commerciaux par pays partenaire. C'est la matière première du projet RICardo. Là encore, la disponibilité de telles statistiques ne va pas de soi. Sur les 24 pays qui établissent des registres de commerce au XVIII^e siècle, 15 fournissent des données de commerce par pays³². La disponibilité de ce type de données s'accroît à partir des années 1830. On soulignera encore le cas particulier de l'Allemagne pour laquelle on ne dispose d'aucune donnée de partenaire commercial avant 1880. Comment sont déterminés les pays de provenance des importations et de destination des exportations ? Il existe, encore de nos jours, différentes méthodes d'enregistrement³³. Le pays de provenance devrait être le pays d'origine première, c'est-à-dire dans lequel la marchandise a été produite ou fabriquée. Mais la méthode du pays de consignment peut être retenue surtout lorsque l'État commerçant est un important centre de transit de marchandises, qui importe pour réexporter (cas du Royaume-Uni de 1905 à 1936). Le pays de provenance peut aussi être le pays de dernière provenance qui n'est en fait qu'un intermédiaire dans l'échange. De même pour définir le pays de destination, on devrait considérer le pays de consommation ou destination finale de la marchandise. Mais il se peut que le partenaire enregistré soit un pays de consignment ou un intermédiaire dans l'opération d'exportation. Les pays enclavés, sans accès à la mer, posent un problème particulier car leurs échanges transitent par les pays frontaliers. Ainsi, ils peuvent « disparaître » des statistiques des pays partenaires au profit d'un pays intermédiaire ou gonfler dans leurs propres statistiques la part des pays intermédiaires. Prenons l'exemple de la Suisse. Elle n'apparaît dans la liste des partenaires commerciaux du Royaume-Uni qu'en 1906, alors que les deux pays commerçaient ensemble depuis bien plus longtemps comme

²⁹ Par exemple, le *Tableau général du commerce de la France* donne, dès le début du XIX^e siècle, un résumé des exportations et des importations distinguant « Matières animales », « Matières végétales », « Matières minérales », et « Fabrications ».

³⁰ C'est en 1853 que le Congrès international de statistique, réuni à Bruxelles, a exprimé le vœu de créer une nomenclature commune des marchandises. Soixante ans plus tard, 29 pays signent la Convention de Bruxelles qui établit la première classification internationale. Elle sera utilisée jusqu'à la publication, en 1938, de la *Liste minimum de marchandises pour les statistiques du commerce international* par la Société des Nations.

³¹ La quatrième et dernière version de la Classification type pour le commerce international, CTIC, (en anglais, Standard International Trade Classification, SITC) des Nations Unies est entrée en vigueur en 2007.

³² Cf. note 19.

³³ Cf. International Monetary Fund, *A Guide to Direction of Trade Statistics*, 1993, p. 4-5.

en attestent les statistiques suisses qui relèvent le partenaire britannique dès 1885³⁴. L'Allemagne représentait alors environ un quart du commerce de la Suisse, on peut se demander si cette part n'est pas surévaluée.

Les méthodes d'élaboration des statistiques commerciales permettent déjà de comprendre le problème de fiabilité posé par les statistiques de commerce bilatéral. C'est un problème majeur que les auteurs du projet RICardo ont choisi de mettre en valeur.

La divergence des flux miroirs

Au-delà des méthodes de valorisation des flux et d'enregistrement des partenaires commerciaux, il existe plusieurs sources de divergence des flux miroirs. La plus évidente est le système d'évaluation c.i.f./f.o.b.³⁵. La valeur des importations inclut généralement les frais de transport, autres frais et taxes y afférents et frais d'assurance, alors que les exportations sont évaluées *franco de port*, sans inclure l'ensemble de ces coûts. C'est pourquoi la valeur des importations mondiales est supérieure à la valeur des exportations mondiales (voir la vue ['World'](#) sur le site RICardo). Ensuite, deux types de commerce doivent être distingués : le commerce général et le commerce spécial. Le premier comprend, outre les marchandises produites, fabriquées ou consommées dans le pays (commerce spécial), des marchandises importées pour être réexportées parfois après transformation. Les données RICardo, comme celles de toutes les bases de données comparables, privilégient le commerce spécial mais les statistiques historiques ne sont pas toujours suffisamment précises et complètes pour que l'on soit assuré de n'avoir que des données de commerce spécial. Il peut aussi y avoir un décalage temporel entre l'enregistrement d'une exportation par un pays et l'enregistrement de l'importation correspondante dans les statistiques du pays importateur. L'évaluation des flux de commerce nécessite des opérations de conversion et le choix du taux de change est une autre source de divergence entre deux flux similaires. Il y a enfin une source de divergence d'origine humaine que sont les erreurs de saisie et la fraude. On estime que la fraude est d'autant plus probable que les taxes douanières sont élevées. Le XIX^e siècle ayant conduit à une relative libéralisation des échanges, donc à une baisse des taxes que l'importateur ou l'exportateur doit payer à l'entrée ou à la sortie du territoire douanier, il est possible que ce problème de fraude soit devenu secondaire. On ne peut cependant pas l'affirmer, d'autant moins que des études récentes montrent que les fausses déclarations impactent encore aujourd'hui les statistiques commerciales³⁶.

Bien que ce problème de divergence soit connu des experts, les flux miroirs sont souvent utilisés pour compléter des données manquantes dans les tableaux de statistiques commerciales. En effet, il n'est pas rare, surtout dans les statistiques historiques, d'avoir des « trous » dans les séries, de ne pas avoir d'information sur le commerce bilatéral de tel pays avec tel ou tel partenaire, telle ou telle année. On notera que les bases de données du FMI peuvent avoir recours aux flux miroirs en appliquant un pourcentage uniforme de 10 % pour obtenir un flux bilatéral d'exportation à partir d'un flux bilatéral d'importation (pour passer de c.i.f. à f.o.b.)³⁷.

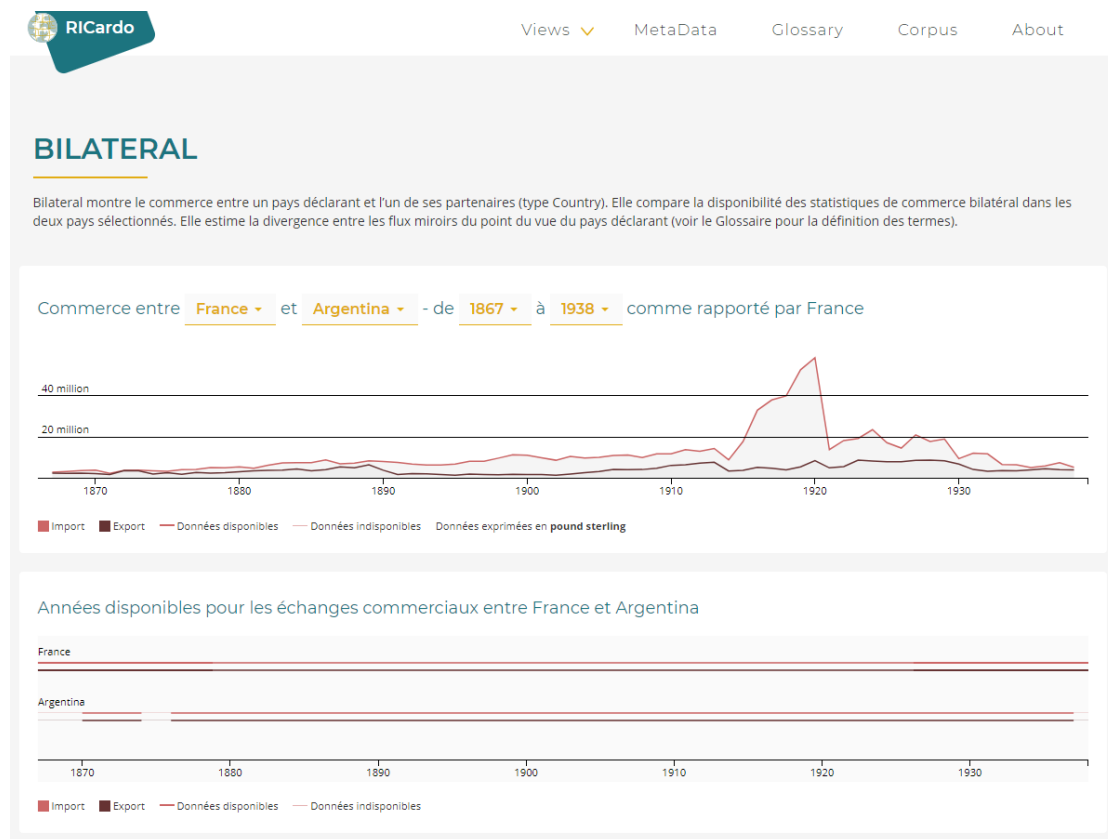
³⁴ On ne dispose pas de document statistique plus ancien.

³⁵ C.i.f. pour *Cost, insurance and freight*. F.o.b. pour *Free on board*.

³⁶ Voir Derek Kellenberg et Arik Levinson, *Misreporting Trade: Tariff Evasion, Corruption, and Auditing Standards*, NBER Working Paper, n° 22593, September 2016.

³⁷ IMF, *op. cit.*, 1993, p. 9-10.

De même, la base de données de Barbieri publiée sur le site de *Correlates of War Project*³⁸ utilise les flux miroirs pour compléter les données manquantes³⁹. L'équipe RICardo a choisi une autre option. Les données manquantes ne sont pas remplacées par les flux miroirs. Il a été décidé d'exploiter la richesse de la base de données pour mettre en évidence le problème de l'imperfection des statistiques commerciales bilatérales. La vue '[Bilateral](#)' proposée sur le site RICardo donne une mesure de l'erreur statistique. L'utilisateur peut sélectionner deux pays partenaires et visualiser instantanément l'importance relative et l'évolution de la divergence entre leurs déclarations commerciales. S'il compare, par exemple, les sources de la France et de l'Argentine, il devra faire deux opérations en sélectionnant d'abord la France puis l'Argentine comme pays rapporteur. Il obtiendra deux graphes différents qui le renseigneront sur la tendance de l'un ou l'autre des deux pays à surévaluer ou sous-évaluer la valeur des échanges.



Vue de la page « Bilatéral France-Argentine »

L'imperfection des statistiques commerciales est un fait que l'on ne peut ignorer. On peut le contourner, ou bien on peut l'observer et l'analyser. Les outils offerts par le projet RICardo peuvent ainsi aider à développer de nouvelles approches en matière de recherche mais aussi d'enseignement.

³⁸ <http://www.correlatesofwar.org/>

³⁹ Katherine Barbieri et Omar Keshk (2016). *Correlates of War Project Trade Data Set Codebook. Version 4.0*, 2016.

Renouveler la recherche et l'enseignement sur l'histoire de la mondialisation et du commerce

RICardo est un outil statistique et numérique. Son élaboration a posé un grand nombre de problèmes méthodologiques qui ont mobilisé les efforts de ses auteurs pendant une dizaine d'années. Ces efforts ont été récompensés puisque la base RICardo est aujourd'hui connue de la communauté scientifique internationale et de plus en plus utilisée pour des tests économétriques sur séries temporelles. Rappelons que c'est l'une des raisons qui ont motivé le lancement du projet. Le passage à la technologie numérique a ouvert de nouvelles perspectives, incitant les auteurs du projet à se poser explicitement la question des innovations que RICardo peut apporter en matière de recherche et d'enseignement.

De nouvelles pistes et méthodes de recherche

Depuis une vingtaine d'années, les principales questions posées par les historiens et/ou économistes intéressés par l'histoire du commerce se concentrent sur le phénomène de mondialisation⁴⁰ et tournent autour de quelques axes qui peuvent être résumés très simplement. La périodisation de la mondialisation : quand a-t-elle commencé, quelles sont les étapes de la mondialisation jusqu'à nos jours (on parle de « première » et de « deuxième » mondialisation) ? Les causes de la mondialisation : le processus d'ouverture aux échanges est-il dû à des innovations technologiques qui réduisent les coûts de transport, à des traités de libre-échange qui réduisent les barrières aux échanges ? Les conséquences de la mondialisation : l'ouverture aux échanges est-elle favorable à la croissance, aux inégalités ? Les formes de la mondialisation : l'intensification des échanges se fait-elle par un volume d'échanges plus important entre mêmes partenaires ou par la création de nouveaux liens commerciaux ? Au vu de ces questions, on peut dire que l'histoire du commerce est surtout abordée d'un point de vue global, à travers le prisme du phénomène de mondialisation.

Travailler à la constitution d'une base de la taille de RICardo, sur plusieurs années, a permis de découvrir les archives commerciales d'un grand nombre de pays de différents continents. Le mot « découvrir » est approprié car il arrive que l'on tombe sur des ouvrages non massicotés qui n'ont jamais été ouverts. L'instinct du chercheur pense aussitôt à la richesse que constituent des documents et données encore largement inexploités. Il s'agirait en l'occurrence de développer des approches locales sur l'histoire commerciale, approches qui peuvent venir contester les points de vue conventionnels ou faire apparaître des points de vue nouveaux⁴¹. D'autre part, la dimension globale de la base, qui ne se concentre pas sur l'Europe et l'Occident, en

⁴⁰ D'après une définition du FMI, le processus de mondialisation commerciale désigne une augmentation de l'interdépendance économique des pays du monde par l'accroissement et la variété des échanges de biens et services.

⁴¹ Pour exemples, voir Béatrice Dedinger, « The Franco-German trade puzzle: an analysis of the economic consequences of the Franco-Prussian war », *Economic History Review*, 2012, 65(3), p. 1029-1054 ; Brian D. Varian, « Anglo-American trade costs during the first era of globalization: the contribution of a bilateral tariff series », *Economic History Review*, 2017 (DOI: 10.1111/ehr.12486). Grâce à l'utilisation de données bilatérales, le premier article montre pour la première fois l'impact de la guerre de 1870 sur la restructuration du commerce franco-allemand ; le second revient sur l'idée admise que le commerce du Royaume-Uni n'a pas été affecté par les droits de douane étrangers entre 1870 et 1913.

fait un outil d'une grande richesse pour explorer de nouvelles visions de la mondialisation au XIX^e siècle comme celle proposée dans le récent ouvrage de deux historiens français⁴². Comme évoqué dans le point précédent, un autre domaine lié à l'histoire du commerce pourrait être approfondi à l'aide de l'outil RICardo, c'est celui qui a trait à la qualité des données commerciales. La vue '*Bilateral*' pourrait être utilisée afin d'établir un classement des pays en fonction de la qualité de leurs statistiques. Mais de nouvelles pistes de recherche sont aussi ouvertes par la dimension technologique du projet. La taille de la base et la création de l'outil numérique permettent l'utilisation de nouvelles techniques de traitement des données. À l'inverse des tests économétriques, dans lesquels les données disparaissent pour se résumer à des coefficients de corrélation et des marges d'erreurs, l'analyse des réseaux est une méthode d'exploration visuelle des données qui peut faire émerger de nouvelles approches de la mondialisation, différentes des problématiques en termes de causalité auxquelles répondent les modèles de régression linéaire⁴³.

Enfin, le projet RICardo a aussi été l'occasion de mettre en œuvre une expérimentation innovante dans la méthode de travail. Grâce à une étroite collaboration interdisciplinaire, une infrastructure de recherche performante a pu être élaborée qui permet le traitement et l'analyse de données historiques hétérogènes et en grand nombre⁴⁴. Des pratiques et outils nouveaux ont été développés pour explorer les données visuellement et de manière interactive, pour créer de nouvelles versions de la base de données et pour tester automatiquement les erreurs et anomalies. En cela, le projet RICardo peut aider d'autres groupes de recherche dans leur pratique. Il est à noter que de nombreux projets, tous centrés sur la reconstitution de bases de données commerciales historiques, sont actuellement développés par des chercheurs de plusieurs pays, et qu'ils s'inspirent et s'enrichissent des méthodes et idées nouvelles apportées par les uns et les autres.

Une autre approche de l'enseignement de la mondialisation

En France, la mondialisation fait partie des programmes d'enseignement à partir de la classe de Première. Les grands thèmes sont abordés : croissance et mondialisation, dynamiques de la mondialisation, mondialisation et finance, mondialisation de la production, régionalisation et mondialisation, l'Union européenne dans l'économie globale, avec une dimension historique jusqu'au début du XIX^e siècle. Le projet RICardo ne vise pas directement à enseigner ou à renouveler l'enseignement de ces grands thèmes sur la mondialisation. La vue '*World*' du site, qui donne l'approche la plus globale de l'évolution des échanges commerciaux, souffre de certains défauts pour les enseignants. Tout d'abord, les valeurs sont exprimées à prix courants. Des ré-estimations à prix constants n'ont pas été tentées car l'objectif premier de RICardo est la compilation d'une base exhaustive de commerce bilatéral. La logique aurait voulu que l'on ré-estime non seulement les séries de commerce total mais aussi les séries de commerce bilatéral, ce qui est une tâche démesurée. Néanmoins, la vue '*World*' intègre les nouvelles séries historiques du commerce mondial estimées par

⁴² Pierre Singaravélou et Sylvain Venayre, *Histoire du Monde au XIX^e siècle*, Fayard, 2017.

⁴³ Voir sur le site du médialab de Sciences Po des exemples d'outils et d'analyses que l'on peut faire autour des réseaux (<http://www.medialab.sciences-po.fr>).

⁴⁴ Paul Girard, Béatrice Dedinger, Donato Ricci, Benjamin Ooghe-Tabanou, Mathieu Jacomy, Guillaume Plique et Gregory Tible, *RICardo Project : Exploring XIX Century International Trade*, Paper presented at Digital Humanities 2016, Krakow, Poland, July 13–15, 2016.

Federico et Tena et renvoie l'utilisateur à leur fichier de données qui propose des séries à prix courants et prix constants⁴⁵. D'autre part, la vue '[World](#)' ne donne pas de vision à long terme de la mondialisation commerciale. Celle-ci est généralement mesurée en calculant un indicateur de degré d'ouverture (exportations/PIB). Une ressource de ce type est proposée aux enseignants sur le site de l'ENS-Lyon⁴⁶. L'intérêt de la vue '[World](#)', d'un point de vue pédagogique, se situe sur un autre plan. Elle permet d'entrer dans la méthodologie de l'élaboration des séries de commerce mondial en montrant quatre estimations différentes⁴⁷. L'élève/étudiant est ainsi amené à s'interroger sur le travail du chercheur et du statisticien : comment est évalué le commerce mondial ? D'où viennent les données ? Pourquoi y a-t-il plusieurs estimations ? Sont-elles très différentes et quel crédit accorder aux estimations du commerce mondial ? Cette vue a aussi un intérêt sur le plan analytique puisqu'elle permet de comparer, en quelques clics, l'évolution de la part des pays dans le commerce mondial.

Si l'on veut développer l'utilisation de RICardo à des fins d'enseignement, il nous faut mettre en avant le fait qu'il met à la disposition du public un corpus de données invitant à une nouvelle approche de l'histoire de la mondialisation et pouvant servir à des travaux pratiques, méthodologiques ou analytiques⁴⁸. Plus encore, RICardo peut aider à faire comprendre à l'élève/étudiant à partir de quel matériau le chercheur écrit l'histoire du commerce et de la mondialisation, et à développer son esprit critique. Pour mieux connaître l'histoire du commerce, d'un point de vue plus local que global, les vues '[World](#)' et '[Country](#)' peuvent être utilisées pour des études historiques sur la part d'un pays dans le commerce mondial et ses relations commerciales bilatérales. La base de taux de change peut servir à illustrer les variations de taux de change sur des périodes tourmentées de l'histoire et fera découvrir à l'utilisateur l'extrême diversité des taux de change. La base des entités est très intéressante pour travailler sur l'histoire politique des États à partir du nom des entités. L'élève retrouvera ainsi que l'Australie ou Commonwealth d'Australie est une création relativement récente et que, avant 1901, existaient six États australiens autonomes (colonies britanniques). La base des entités peut servir à reconstituer le commerce d'une métropole avec ses colonies. Les relevés statistiques du XIX^e siècle mentionnent souvent les ports de commerce plutôt que les pays. C'est un moyen de faire connaître les routes commerciales aux XIX^e et début XX^e siècles. Enfin, la base des sources offre un grand choix d'utilisations possibles que l'utilisateur peut apprécier dans la vue '[Metadata](#)'. Cette vue recense l'ensemble des sources consultées pour créer la base de données commerciales et elle permet de voir quelle source a été utilisée pour chaque flux de commerce. Toutes les sources sont classées suivant qu'elles sont primaires (tableaux statistiques d'origine), secondaires (compilations de sources primaires) ou des estimations (travaux de recherche effectués à partir de sources primaires ou secondaires). C'est une base déjà extrêmement utile pour le chercheur, mais elle a aussi un potentiel pédagogique pour apprendre aux élèves/étudiants à maîtriser les outils et les méthodes de l'historien. Qu'est-ce qu'une source de commerce ? Quelle est la qualité des différentes sources ? Que trouve-t-on

⁴⁵ Giovanni Federico et Antonio Tena-Junguito, *op. cit.*, 2016.

⁴⁶ <http://ses.ens-lyon.fr/ressources/stats-a-la-une/levolution-de-louverture-commerciale-de-1827-a-2014> [lien consulté le 15/02/2018].

⁴⁷ Voir leur définition sous l'onglet '[Glossaire](#)' du site.

⁴⁸ Les travaux de Thomas Piketty ont beaucoup aidé à éveiller l'intérêt du grand public pour des études historiques fondées sur des séries statistiques longues.

dans les sources sur le commerce ? Pour parfaire cette base, des photos numérisées des tableaux statistiques seront également mises en ligne.

Ce n'est pas un hasard si le projet RICardo s'est développé à une époque où le phénomène de *big data* est devenu un grand défi informatique. La conception d'un tel projet n'aurait pas été possible il y a seulement vingt ans et les résultats obtenus n'avaient pas été envisagés il y a seulement dix ans. L'expérience RICardo est un exemple concret de l'adaptation progressive des sciences humaines et sociales à la révolution numérique. Parti d'un objectif de construction de base de données avec des moyens qui en limitaient la portée, le projet RICardo a connu une réelle transmutation pour devenir un outil de travail et de diffusion de la connaissance potentiellement grand public. Il peut encore être amélioré. Trois grands axes sont actuellement envisagés pour continuer à innover et développer les performances de l'outil : 1. créer sur le site une visualisation cartographique des relations commerciales bilatérales ; 2. introduire dans la base des entités une information nouvelle sur le statut politique des entités et son évolution dans le temps. Cela permettrait d'élargir considérablement le potentiel d'utilisation de RICardo en facilitant l'accès à des problématiques telles que l'intégration commerciale, le commerce des empires coloniaux ou des unions douanières ; 3. le dernier axe est de prolonger jusqu'à nos jours la période couverte par la base en incluant la base de données bilatérales du FMI. Il ne s'agirait pas d'apporter des données nouvelles mais un service étendu à l'utilisateur qui aurait à sa disposition des séries de commerce bilatéral et total les plus complètes possibles.

Le progrès technique n'est cependant pas une condition suffisante pour assurer le succès d'une telle entreprise. Comme il a été souligné dans cet article, ce projet n'aurait jamais pu aboutir sans la participation active de nombreux chercheurs et étudiants, sans les encouragements et les conseils de collègues français et étrangers, et sans la coordination et la détermination des meneurs du projet.