

ÉCOLE NATIONALE DES CHARTES

Jean-Damien Généro

Licencié ès histoire

Diplômé de Master d'histoire médiévale

VALORISER LE TRAITEMENT AUTOMATIQUE DES DONNÉES :

Le cas des Ouvriers des deux mondes

Mémoire pour le diplôme de Master
« Technologies numériques appliquées à l'histoire »

2020

Résumé

Résumé d'une trentaine de lignes à placer en tête du mémoire, accompagné d'une dizaine de mots-clés destinés à décrire le mémoire et des informations bibliographiques nécessaires pour le citer. Ce résumé et ces mots-clés sont destinés à compléter la notice bibliographique du mémoire dans la future bibliothèque numérique des mémoires.

N.B. : ne pas dépasser une page pour le tout.

Mots-clés : Liste d'une dizaine de mots-clés, séparés par des points-virgules.

Informations bibliographiques : Jean-Damien Généro, *Valoriser le traitement automatique des données : Le cas des Ouvriers des deux mondes*, mémoire de Master « Technologies numériques appliquées à l'histoire », dir. A. Chagué et V. Jolivet, École nationale des chartes, 2020.

Liste des sigles et abréviations

Institutions

- ALMAaCH : *Automatic Language Modelling and Analysis & Computational Humanities*
 - ANR : Agence nationale de la recherche
 - CNRS : Centre National de Recherche Scientifique
 - EHESS : École des Hautes Études en Sciences Sociales
 - LARHRA : Laboratoire de Recherche Historique Rhône-Alpe
-

Langages informatique

- JSON : *JavaScript Object Notation*
- TEI : *Text Encoding Initiative*
- XML : *eXtensible Markup Language*
- XSLT : *eXtensible Stylesheet Language Transformations*

Préambule : un stage confiné ?

Le stage qui a donné lieu au présent mémoire n'a pas commencé sous les meilleurs auspices. Des difficultés administratives et procédurales, puis le grand confinement et les difficultés plus grandes encore qui en ont résulté, ont failli avoir raison de lui. Il a fallu se battre pour que l'Université de Paris admette que la fermeture des établissements d'enseignement annoncée par le président de la République le 12 mars 2020, puis les mesures de confinement à partir du 17 mars, ne signifiaient en aucun cas une suspension de l'action de son administration et un report automatique et sans appel des procédures en court, dont celle concernant les stages. Je tiens donc en premier lieu à remercier Mesdames Alix Chagué et Manuela Martini, Messieurs Julien Cassefières et Thibault Clérice, qui m'ont apporté un tant soit peu de soutien et d'aide dans cette procédure extrêmement pénible, absolument anormale et pour laquelle j'ai dépensé une quantité d'énergie démesurée.

Contactée, la présidence de l'Université de Paris mit fin au marasme, apposa les cachets requis et présenta ses excuses ; le stage pu ainsi commencer, avec une semaine de retard sur la date prévue.

Intégrer une équipe d'ingénierie et de recherche tout en étant confiné chez soi n'est pas chose aisée. Les relations humaines que cet exercice suppose en temps normal ont été réduites au minimum et sont passées par des courriels, des échanges sur l'espace de discussion instantanée d'INRIA et, principalement, des visioconférences sur *Zoom*.

Inconnue, immédiatement décevante et parfois pénible, cette façon de travailler a mis un certain temps à s'imposer pour moi et les premières semaines ont été difficiles. Début juin, c'est-à-dire à la moitié du stage, Alix Chagué me demanda de lui transmettre un bilan des compétences que j'avais acquises jusqu'ici. Elles étaient nombreuses du point de vue des savoir-faire, car, bon gré mal gré, je n'avais eu d'autre choix que de m'adapter à ce travail à distance et à remplir les missions qui m'avaient été assignées. Ce faisant, un des objectifs que je m'étais fixé pour ce stage fut très vite satisfait : mes savoir-faire se sont développés par l'apprentissage et la maîtrise de nouvelles techniques.

Quid des savoir-être ? J'avais listé plusieurs points dans le bilan de mi-stage, sur un ton légèrement humoristique, et notamment saluer, faire une présentation et suivre une réunion en visioconférence. Un véritable savoir-être se cachait derrière cet humour désabusé : ce stage m'a appris le travail à distance. Cela suppose beaucoup de choses. Outre une organisation et une discipline peut-être plus importantes que pour un travail de bureau, il s'agit d'être seul la majorité du temps. De ne pas avoir de collègue avec qui l'on peut échanger et progresser dans la résolution d'un problème. En un mot, il s'agit de travailler plus avec soi-même qu'avec les autres et donc de se *débrouiller*.

Introduction

Les *Ouvriers des deux mondes* au sein de l'ANR *Time Us*

« Reconstituer les rémunérations et les budgets temps des travailleuses et des travailleurs du textile dans quatre villes industrielles françaises (Lille, Paris, Lyon, Marseille) dans une perspective européenne et de longue durée »¹ est l'objectif du programme ANR *Time Us*. Son intitulé exact, « Rémunérations et usages du temps des femmes et des hommes en France de la fin du XVII^e siècle au début du XX^e siècle »² met en évidence sa double articulation autour d'un temps long qui court depuis les premières manufactures de l'époque Moderne jusqu'à la fin de la révolution industrielle au tournant des XIX^e et XX^e siècles, et du temps individuel et quotidien qui est celui d'un ouvrier ou d'une ouvrière au sein de ces grands mouvements historiques. Les relations entre la rémunération et le temps sont au cœur des questionnements du programme. Cette attention portée à l'usage du temps se traduit dans la dénomination courante du programme, « *Time Us* », abréviation de l'anglais « *time usage* ».

Initialement financé sur une durée de trente-six mois (janvier 2017-janvier 2020), le programme ANR, coordonné par le professeur Manuela Martini de l'Université Lumière Lyon 2, est porté par une équipe pluri-institutionnelle. Celle-ci est composée de quatre unités mixtes de recherche (UMR) formées d'universitaires et de chercheurs du Centre National de la Recherche Scientifique³, d'une équipe d'accueil (EA)⁴ de l'Université de Paris⁵ et de l'équipe projet ALMAnaCH⁶ de l'Institut national de recherche en informatique et en automatique (Inria⁷ Paris). Le stage qui a donné lieu au présent mémoire était placé sous la responsabilité administrative du laboratoire ICT de l'Université de Paris et a été effectué sous le tutorat scientifique et professionnel d'Inria, en la personne d'Alix Chagué, ingénierie de recherche et de développement au sein de l'équipe ALMAnaCH.

Time Us se concentre principalement sur les femmes et plus encore sur les ouvrières du textile, industrie dans laquelle « elles sont présentes dans toutes les phases du proces-

1. Présentation du programme sur le site de l'ANR (<https://anr.fr/Projet-ANR-16-CE26-0018>, consulté le 20 août 2020).

2. Présentation du programme sur le site du LARHRA (<http://larhra.ish-lyon.cnrs.fr/anr-time-us>, consulté le 20 août 2020).

3. Le Laboratoire de recherches historiques Rhône-Alpes (LARHRA, Lyon 2), le laboratoire Temps, Espaces, Langages, Europe Méridionale-Méditerranée (TELEMMe, Université d'Aix-Marseille), l'Institut de Recherches Historiques du Septentrion (Université de Lille) et le Centre Maurice-Halbwachs (CMH, EHESS et ENS Paris).

4. Le laboratoire Identités, Cultures et Territoires (ICT).

5. Université formée par décret n° 2019-209 du 20 mars 2019 et regroupant les établissements Paris V-Descartes et Paris VII-Diderot.

6. *Automatic Language Modelling and Analysis & Computational Humanities*.

7. « Inria » est depuis 2011 une marque; aussi l'élision par le biais d'un « l' » liminaire n'est-elle plus requise dans son écriture courante (« Rappel : désormais *L'INRIA* s'écrit *Inria* », <https://web.archive.org/web/20111113064215/https://www.inria.fr/institut/inria-en-bref/charte-logo-inria/charte>, consulté le 20 août 2020).

sus productif »⁸. Le programme tend à combler le biais des genres dans l'historiographie du travail industriel en réalisant une opération de collecte et de traitement de la documentation manuscrite et imprimée relative à l'emploi et aux activités quotidiennes des femmes⁹. Ainsi, le moteur de *Time Us* est moins la production d'une réflexion scientifique autour du travail des femmes que la constitution d'un corpus documentaire sériel et prêt à être exploité par des chercheurs d'horizons multiples. La pluridisciplinarité est un aspect majeur du programme, qui souhaite une utilisation de ses données dans un maximum de champs de recherche des sciences humaines et sociales, notamment « en histoire économique et sociale, en histoire de la famille et du genre, en histoire des conflits du travail et de la culture des classes populaires »¹⁰.

La documentation est essentiellement constituée de documents manuscrits conservés par la Bibliothèque municipale de Lyon et les dépôts d'archives départementaux et municipaux lillois, parisien, lyonnais et marseillais. Très diverse, elle est issue d'organes officiels, à l'instar des conseils prud'homaux, des chambres de commerce ou encore des tribunaux de commerce, mais aussi des archives de personnes physiques ou morales de droit privé, telles que celles du tisseur lyonnais Pierre Charnier (1795-1857)¹¹, ou les registres comptables (1817-1821) et le dossier de faillite (1821-1822) de la filature parisienne Dupuis-Drouet¹².

En sus de cette documentation manuscrite, le programme *Time Us* s'appuie sur trois grands corpus d'imprimés. Le premier se compose de neuf titres de la presse ouvrière lyonnaise, entièrement numérisés sur le site *Numelyo*, espace digital de la Bibliothèque municipale de Lyon¹³. L'intérêt du programme pour ce corpus porte sur les nombreux compte rendus d'audience du Conseil des prud'hommes qui s'y trouvent, mais aussi pour les reproductions ou extraits de discours, les lettres ou encore les analyses économiques et sociales relayés par ces journaux. Quatre des titres sont publiés dans la première moitié des années 1830 (*L'Écho de la fabrique*, *L'Écho des travailleurs*, *L'Indicateur* et *La Tribune proléttaire*), les cinq restant l'étant dans les années 1840 (*L'Écho des ouvriers*, *L'Écho de la Fabrique de 1841*, *L'Écho de l'industrie*, *L'Avenir*, *La Tribune Lyonnaise*)¹⁴.

8. Alix Chagué, Manuela Martini, Victoria Le Fournier et Éric Villemonte de la Clergerie, *Deux siècles de sources disparates sur l'industrie textile en France : comment automatiser les traitements d'un corpus non-uniforme ?*, 2020, p. 1.

9. *Ibid.*

10. *Ibid.*, p. 2.

11. Les papiers du canut Pierre Charnier font partie du fonds « Fernand Rude » de la Bibliothèque municipale de Lyon, nommé d'après l'historien qui les possédait avant leur versement (<https://www.bm-lyon.fr/collections-patrimoniales-et-specialisees/explorer-les-collections/article/fernand-rude>, consulté le 20 août 2020).

12. Archives de Paris, D 12U1, n° 375-376 (http://archives.paris.fr/arkotheque/inventaires/ead_ir_consult.php?a=4&ref=FRAD075_000727, consulté le 20 août 2020).

13. Le corpus se trouve à l'adresse <https://collections.bm-lyon.fr/PER003> (consulté le 20 août 2020).

14. Présentation de ce corpus sur le wiki du programme : http://timeusage.paris.inria.fr/mediawiki/index.php/Documentation_régionale_-_Presse_lyonnaise (consulté le 20 août 2020).

Un second corpus d'imprimés est constitué de monographies collectées sur le site *Gallica* de la Bibliothèque nationale de France au format PDF. Très divers, on y trouve à la fois *L'ouvrière* de Jules Simon (1861), un *Dictionnaire général des tissus anciens et modernes* (1859-1863) ou encore un *Traité complet sur la fabrication des étoffes de soie* (1859)¹⁵. Il s'agit d'une base complémentaire au projet, qui doit faciliter la contextualisation de la base archivistique principale, notamment en permettant de suivre l'évolution du vocabulaire afférent au textile sur la période étudiée.

Le troisième et dernier corpus d'imprimés est composé des treize volumes de la série des *Ouvriers des deux mondes*.

*

Initiées par le sociologue Frédéric Le Play (1806-1882), les *Ouvriers des deux mondes* sont des enquêtes sociologiques conduites par les membres de la Société internationale des études pratiques d'économie sociale¹⁶ de 1857 à 1928. Répartie en trois séries comptant un total de cent vingt-six monographies¹⁷, il s'agit de la deuxième entreprise d'étude empirique de Le Play, après celle des *Ouvriers européens* dont la première édition a lieu en 1855¹⁸. Les enquêtes sont couramment désignées sous le vocable de « monographies de famille » ou plus simplement « monographies ».

Dans sa préface du numéro spécial de la revue *Les Études sociales* consacré aux monographies leplaysiennes, Alain Chenu relève la prégnance de l'assimilation des monographies leplaysiennes à des « mines » dans lesquelles les chercheurs peuvent « [puiser] » à leur guise, tant les sujets abordés et les étendues géographiques traitées sont nombreux¹⁹. *Les Ouvriers des deux mondes*, dont le titre fait écho à *La revue des deux mondes* fondée en 1829, présentent en effet une succession de « trajectoires et récits de vie de familles ouvrières »²⁰ établies de part et d'autre de la Méditerranée. Se succèdent ainsi une enquête consacrée à un charpentier de Paris²¹, à un métayer de la banlieue de Florence²² et à un

15. Liste complète sur le wiki du programme : http://timeusage.paris.inria.fr/mediawiki/index.php/Aperçu_des_états#Imprimés_divers (consulté le 20 août 2020).

16. Actuelle Société d'économie et de sciences sociales, désormais abrégée en SESS.

17. Anthony Lorry, « Les monographies des *Ouvriers européens* (1855, 1877-1879) et des *Ouvriers des deux mondes* (1857-1930). Inventaire et classification », *Les Études sociales, Les monographies de famille de l'École de Le Play*-n° 131-132 (2000), p. 93-181, p. 95.

18. *Ibid.*

19. Alain Chenu, « Préface », *Les Études sociales, Les monographies de famille de l'École de Le Play*-n° 131-132 (2000), p. 5-9, p. 5.

20. Stéphane Baciocchi et Jérôme David, « Trajectoires et récits de vie. Singulariser les expériences sociales », *Les Études sociales, Frédéric Le Play : anthologie et correspondance*, n° 142-143-144 (II-2005-2006), p. 193-194, URL : https://www.academia.edu/5925885/Fr%C3%A9d%C3%A9ric_Le_Play_%C3%A9mants_d'%C3%A9pist%C3%A9mologie_et_de_science_sociale (visité le 11/08/2020), p. 193.

21. Frédéric Le Play et Adolphe Focillon, « Charpentier de Paris (Seine - France), de la Corporation des compagnons du Devoir », dans *Les Ouvriers des deux mondes*, Paris, 1857 [janvier 1858] (série 1 (1)), t. 1, chap. 1, [27]-68.

22. Ulbadino Péruzzi, « Métayer de la banlieue de Florence (Grand-Duché de Toscane) », dans *Les Ouvriers des deux mondes*, Paris, 1857 [janvier 1858] (série 1 (1)), chap. 5, p. 221-262.

menuisier-charpentier de Tanger²³.

Les *Ouvriers des deux mondes* présentent un double intérêt pour *Time Us*. D'une part, leur attention dirigée de manière exclusive envers les familles ouvrières en font un matériau privilégié pour ce programme, d'autant que les monographies se focalisent sur le budget et son usage²⁴. Aucun individu de la cellule familiale n'est ignoré : l'ouvrier — il s'agit le plus souvent du père, qui peut être accompagné de ses frères ou de ses fils —, ses descendants, ses ascendants, sa femme évidemment, mais aussi les domestiques²⁵ et parfois les esclaves²⁶. Ainsi, si seulement quatorze enquêtes ont pour sujet des familles travaillant dans l'industrie du textile, l'ensemble reste intéressant à l'échelle du programme ANR dans la mesure où chaque enquête s'attache à établir le budget affecté aux matières textiles et à son utilisation par la famille enquêtée. L'angle de lecture adopté par *Time Us* concerne donc la totalité du texte pour quatorze monographies, et se focalise sur les paragraphes et tableaux relatifs aux individus et à leurs budgets pour les cent douze restantes.

D'autre part, la reproduction systématique d'une même structure logique dans chaque monographie permet d'envisager un traitement informatique de l'ensemble du corpus et, à travers celui-ci, la mise à disposition des textes structurés dans des fichiers XML. L'idée des « mines » leplaysiennes est en effet inexacte, et Alain Chenu dénonce le caractère restrictif de cette image : les monographies ne sont pas une succession d'enquêtes indépendantes. Le Play et la SESS ont conservé une même « grille d'observation » depuis la première (1856) jusqu'à la dernière enquête (1928), construisant *de facto* un « système dont les éléments prennent sens les uns par rapport aux autres »²⁷.

Time Us a souhaité transcrire automatiquement les volumes des *Ouvriers des deux mondes* à partir de leurs numérisations, avant d'implémenter la structure logique d'origine au sein du texte brut obtenu. Le but visé est de mettre à disposition de la communauté scientifique ces textes afin qu'elle puisse y porter un regard nouveau à l'aide des outils numériques ; il s'agit de faire entrer ce « matériau exceptionnel sur les sociétés des cinq continents »²⁸ dans le champ des humanités numériques.

*

Les rapports entre le programme ANR *Time Us* et le Master « Technologies numériques appliquées à l'histoire » de l'École nationale des chartes remontent à 2018. Deux

23. Narcisse Cotte, « Menuisier-charpentier (Nedjar) de Tanger (Province de Tanger - Maroc) », dans *Les Ouvriers des deux mondes*, Paris, 1858 (série 1 (2)), chap. 12, p. 105-144.

24. Le budget est « à la fois la méthode et le résultat » des monographies : Fabien Cardoni, « Aux sources du budget domestique selon Le Play », *Les Études sociales*-1 (2012), p. 11-46, URL : <https://doi.org/10.3917/etsoc.155.0011>, p. 11

25. Ernest Delbet, « Paysans en communauté et en polygamie de Bousrah (Esky Cham), dans le pays de Haouran (Syrie - Empire Ottoman) », dans *Les Ouvriers des deux mondes*. Paris, 1858 (série 1 (2)), chap. 18, p. 363-446.

26. Narcisse Cotte, « Menuisier-charpentier... », *op. cit.*

27. Alain Chenu, *op. cit.*, p. 5.

28. Antoine Savoye, « Editorial », *Les Études sociales*, *Les monographies de famille de l'École de Le Play*-n° 131-132 (2000), p. 3-4.

étudiantes, Alix Chagué (prom. 2018) et Victoria Le Fournier (prom. 2019) y ont effectué leur stage de fin d'étude. Alix Chagué a travaillé sur la collecte et le traitement de la documentation²⁹, tandis que Victoria Le Fournier s'est intéressée à la structuration de cette documentation en prenant pour exemple la juridiction prud'homale³⁰. Leurs mémoires correspondent à deux phases successives du projet, et la présente étude entend en décrire la phase finale correspondant à la valorisation des fichiers obtenus. Si cet aspect l'inscrit dans la suite des précédents, il n'en reste pas moins fondamentalement différent.

En premier lieu, il porte une attention exclusive à un seul corpus du programme, les *Ouvriers des deux mondes*. Alix Chagué balayait pour sa part l'ensemble de la documentation et Victoria Le Fournier s'appuyait quant à elle sur une de ses composantes majeures et transversales. Les documents afférents aux conseils des prud'hommes se trouvent en effet à la fois dans les archives propres à ces institutions et, pour la juridiction lyonnaise, sous forme de compte rendus d'audience publiés dans la presse. Le présent mémoire ne cherche ni à replacer les *Ouvriers des deux mondes* au sein de la documentation globale du programme, ni à décrire des opérations de traitement et de mise en forme de ce corpus.

En effet, au commencement du stage ayant donné lieu à ce mémoire en avril 2020, les *Ouvriers des deux mondes* étaient déjà entièrement mis en forme et structurés au format XML-TEI. Ainsi — c'est la deuxième différence d'avec nos prédecesseures — nous sommes intervenus dans la phase de post-traitement, c'est-à-dire que l'essentiel du stage a été consacré au contrôle qualité des fichiers, à la réparation des erreurs issues du traitement automatique et à une réflexion sur la valorisation à mener des données ainsi produites.

Plusieurs objectifs ont été poursuivis durant ce stage, et les questionnements qu'ils portent sous-tendent le propos que nous allons tenir entre ces pages. Il faut tout d'abord rappeler que le mémoire du Master professionnel « Technologies numériques appliquées à l'histoire » est un exercice à mi-chemin entre le mémoire de recherche et le rapport de stage. Il ne s'agit pas d'établir un compte rendu des missions reçues et des actions menées, ni d'exposer une réflexion scientifique sur un sujet délimité, mais de mettre en perspective de manière critique l'intervention du stagiaire en la replaçant au sein des dynamiques scientifiques propres au projet et à l'équipe où le stage a été effectué.

Dans cet optique, notre intervention avait un premier objectif technique et professionnel. Il s'agissait d'analyser et de reprendre des fichiers XML-TEI afin de corriger leurs erreurs structurelles afin d'amorcer leur valorisation. Pour mener à bien cette opération, nous avions la possibilité de recourir à des techniques automatiques ou semi-automatiques

29. A. Chagué, *Constituer un corpus pour la fouille de texte — de la transcription des documents d'archives à l'annotation, Exploration d'une méthodologie par l'ANR Time Us*, mémoire du Master « Technologies numériques appliquées à l'histoire », École nationale des chartes, 2018, URL : <https://github.com/alix-tz/M2TNAH-memoireDeStage>.

30. V. Le Fournier, *Étude de la structuration automatique et de l'éditorialisation d'un corpus hétérogène, L'exemple des sources du conseil des prud'hommes pour le textile au XIX^e siècle du projet Time Us*, mémoire du Master « Technologies numériques appliquées à l'histoire », École nationale des chartes, 2019, URL : <https://github.com/Victorialf/M2TNAH-memoireDeStage>.

en appliquant des scripts Python, ou bien d'effectuer les corrections en manuel. Le second objectif a résulté de cette façon de faire et interrogeait la valeur ajoutée de l'automatisation pour l'édition numérique d'une documentation imprimée dans le cadre d'un programme de recherche.

Nous tiendrons dans ce mémoire un propos en trois parties. Pour commencer, nous présenterons les différents états du corpus des *Ouvriers des deux mondes*, qui est à la fois imprimé, numérisé et structuré en fichiers informatiques. Cela sera pour nous l'occasion de décrire la structure logique mise en place par Le Play et de faire la part entre la volonté de continuité de ses successeurs dans la réutilisation de celle-ci et les innovations qu'ils ont pu mettre en place au fil du temps. Nous conclurons cette première partie par la description des opérations qui ont permis d'obtenir les fichiers des monographies, qui nous furent transmis sous la forme d'un dépôt au sein de l'espace GitLab d'INRIA.

La seconde partie abordera les reprises que nous avons effectuées, en commençant par définir la méthode de travail suivie et notamment les outils de développements et les espaces d'échange mis en place pour la gestion du projet. Après avoir présenté une typologie des erreurs identifiées à l'issue de l'analyse des documents, nous exposerons nos interventions tant manuelles qu'automatiques en procédant du général au particulier, c'est-à-dire du niveau global du corpus à celui, plus fin, des fichiers.

Nous réfléchirons enfin aux différentes manières de valoriser ces fichiers dans la troisième partie, en insistant sur trois voies possibles. La première consiste à déterminer l'opportunité de conserver un lien entre le texte de la page et son image d'origine, ainsi que la nature que celui-ci pourrait prendre (stockage local, utilisation des ressources de la plate-forme d'hébergement des numérisations, recours à un protocole IIIF³¹). La seconde concerne l'indexation des individus enquêtés et la place que pourrait tenir l'automatisation dans cette opération. La dernière porte enfin sur la possibilité de corriger les transcriptions pour que les fichiers soient exploitables non seulement au niveau des données qu'ils contiennent, mais aussi au niveau du texte en lui-même, c'est-à-dire de permettre et à la machine et à l'humain de les utiliser³².

31. Abréviation d'*International Image Interoperability Framework*, généralement prononcée « triple I F » en français.

32. Ce mémoire est entièrement rédigé avec le langage LATEX. Les fichiers qui le constituent sont disponibles sur le GitHub de l'auteur, consultables et téléchargeables à cette adresse : <https://github.com/jeandamien-genero/Memoire-TNAH>.

Bibliographie

Bibliographie

Bibliographie générale

- BACIOCCHI (Stéphane) et DAVID (Jérôme), « IV. Ramifications épistémologiques. Décrire, définir, évaluer, comparer », *Les Études sociales, Frédéric Le Play : anthologie et correspondance*, n° 142-143-144 (II-2005-2006), p. 87-90, URL : https://www.academia.edu/5925885/Fr%C3%A9d%C3%A9ric_Le_Play_%C3%A9t%C3%A9ments_d%C3%A9pist%C3%A9mologie_et_de_science_sociale (visité le 11/08/2020).
- « Trajectoires et récits de vie. Singulariser les expériences sociales », *Les Études sociales, Frédéric Le Play : anthologie et correspondance*, n° 142-143-144 (II-2005-2006), p. 193-194, URL : https://www.academia.edu/5925885/Fr%C3%A9d%C3%A9ric_Le_Play_%C3%A9t%C3%A9ments_d%C3%A9pist%C3%A9mologie_et_de_science_sociale (visité le 11/08/2020).
- CARDONI (Fabien), « Aux sources du budget domestique selon Le Play », *Les Études sociales*–1 (2012), p. 11-46, URL : <https://doi.org/10.3917/etsoc.155.0011>.
- CHAGUÉ (Alix), *Constituer un corpus pour la fouille de texte — de la transcription des documents d'archives à l'annotation, Exploration d'une méthodologie par l'ANR Time Us*, mémoire du Master « Technologies numériques appliquées à l'histoire », École nationale des chartes, 2018, URL : <https://github.com/alix-tz/M2TNAH-memoireDeStage>.
- CHAGUÉ (Alix), MARTINI (Manuela), LE FOURNER (Victoria) et VILLEMONTE DE LA CLERGERIE (Éric), *Deux siècles de sources disparates sur l'industrie textile en France : comment automatiser les traitements d'un corpus non-uniforme ?*, 2020.
- CHENU (Alain), « Préface », *Les Études sociales, Les monographies de famille de l'École de Le Play*–n° 131-132 (2000), p. 5-9.
- LE FOURNER (Victoria), *Étude de la structuration automatique et de l'éditorialisation d'un corpus hétérogène, L'exemple des sources du conseil des prud'hommes pour le textile au XIX^e siècle du projet Time Us*, mémoire du Master « Technologies numériques appliquées à l'histoire », École nationale des chartes, 2019, URL : <https://github.com/Victorialf/M2TNAH-memoireDeStage>.
- LORRY (Anthony), « Les monographies des *Ouvriers européens* (1855, 1877-1879) et des *Ouvriers des deux mondes* (1857-1930). Inventaire et classification », *Les Études*

- sociales, Les monographies de famille de l'École de Le Play*–n° 131-132 (2000), p. 93-181.
- SAVOYE (Antoine), « Éditorial », *Les Études sociales, Les monographies de famille de l'École de Le Play*–n° 131-132 (2000), p. 3-4.
- « La monographie sociologique : jalons pour son histoire (1855-1974) », *Les Études sociales, Les monographies de famille de l'École de Le Play*–n° 131-132 (2000), p. 11-46.

Les Ouvriers des deux mondes

- BIGOT (Maximilien) et ESCARD (François), « Paysans corses en communauté, porchers-bergers des montagnes de Bastelica », dans *Les Ouvriers des deux mondes*, Paris, 1890 (série 2e (2)), p. 433-524.
- COTTE (Narcisse), « Menuisier-charpentier (Nedjar) de Tanger (Province de Tanger - Maroc) », dans *Les Ouvriers des deux mondes*, Paris, 1858 (série 1 (2)), chap. 12, p. 105-144.
- DELAIRE (Alexis), « Paysan-paludier du Bourg de Batz (Loire-Inférieure - France) », dans *Les Ouvriers des deux mondes*. Paris, 1887 (Fascicule 1 (01)), p. 1-56.
- DELBET (Ernest), « Paysans en communauté et en polygamie de Bousrah (Esky Cham), dans le pays de Haouran (Syrie - Empire Ottoman) », dans *Les Ouvriers des deux mondes*. Paris, 1858 (série 1 (2)), chap. 18, p. 363-446.
- LE PLAY (Frédéric), « Introduction. III, Appréciation des deux procédés communément employés pour observer les faits sociaux : supériorité des enquêtes directes sur les recherches scientifiques », dans *Les Ouvriers européens*, Paris, 1855, p. 9-12, URL : <https://gallica.bnf.fr/ark:/12148/bpt6k1057844n/f17>.
- « Instruction sur la méthode d'observation dite des monographies de famille, propre à l'ouvrage intitulé *Les ouvriers européens* », dans *Les Ouvriers des deux mondes*, Paris, 1862 (série 1 (4)), p. 15-32, URL : <https://archive.org/details/lesouvriersdesde04sociuoft/page/14/>.
- LE PLAY (Frédéric) et FOCILLON (Adolphe), « Charpentier de Paris (Seine - France), de la Corporation des compagnons du Devoir », dans *Les Ouvriers des deux mondes*, Paris, 1857 [janvier 1858] (série 1 (1)), t. 1, chap. 1, [27]-68.
- PÉRUZZI (Ulbadino), « Métayer de la banlieue de Florence (Grand-Duché de Toscane) », dans *Les Ouvriers des deux mondes*, Paris, 1857 [janvier 1858] (série 1 (1)), chap. 5, p. 221-262.
- SOCIALE (Société internationale des études pratiques d'économie), « Avertissement », dans *Les Ouvriers des deux mondes*, Paris, 1875 (série 1 (5)), p. I-III, URL : <https://archive.org/details/lesouvriersdesde05sociuoft/page/n9>.

« Avertissement », dans *Les Ouvriers des deux mondes*. éd. Société internationale des études pratiques d'économie sociale, Paris, 1887, p. I-VIII.

Première partie

Un corpus déjà structuré

Le corpus sur lequel nous avons travaillé est un ensemble de fichiers XML-TEI. Ces derniers résultent d'une version numérisée des *Ouvriers des deux mondes* réalisée et hébergée par le site *Internet Archive*, elle-même issue du traitement des volumes physiques de l'Université de Toronto. Dans cette première partie, nous allons décrire ces différents états des *Ouvriers des deux mondes* et commenter les opérations qui ont mené à leur structuration, tout en interrogeant les raisons ayant conduit le programme *Time Us* à les privilégier par rapport à d'autre version similaires.

Chapitre 1

Un corpus d'imprimés

1.1 Une publication sur une soixantaine d'années

La publication des trois séries des *Ouvriers des deux mondes* s'échelonne sur soixante-treize années (1857-1930). Elle n'est véritablement continue que sur trois périodes, de 1857 à 1862, puis de 1885 à 1913 et enfin de 1928 à 1930. L'interruption entre 1913 et 1928 est une conséquence de la première Guerre mondiale et des difficultés budgétaires auxquelles la SESS doit faire face après celle-ci. La première interruption, entre 1862 et 1885, est quant à elle due à un recalibrage des objectifs de la SESS, qui considère que « continuer à recueillir des faits sans essayer d'en faire sortir aucune conclusion (...) c'eût été (...) une preuve d'impuissance et de stérilité »¹. Elle se concentre dès lors sur la parution de son *Bulletin* jusqu'au début des années 1880, où « le besoin de conclusions pratiques était amplement satisfait » et « reprendre l'œuvre des monographies de familles » devenait possible².

Une tentative de reprise avait déjà été effectuée en 1875 où « un premier fascicule composé de trois monographies » avait paru³. Deux autres paraissent en 1883 et 1884, et les trois sont finalement rassemblés en 1885 dans le cinquième volume de la première série⁴. Ce volume apparaît ainsi à bien des égards comme un moment charnière dans l'histoire de la publication des *Ouvriers des deux mondes*. En effet, les monographies suivantes des deuxième (1887-1899) et troisième série (1904-1930) sont publiées sous forme de fascicules trimestriels qui sont ensuite reliés en volumes, là où les quatre premiers tomes étaient directement parus en volumes⁵. Des éléments de paratexte sont systématiquement fournis aux relieurs, notamment une page de titre, une introduction au volume, un index,

1. Société internationale des études pratiques d'économie sociale, « Avertissement », dans *Les Ouvriers des deux mondes*, Paris, 1875 (série 1 (5)), p. I-III, URL : <https://archive.org/details/lesouvriersdesde05sociuoft/page/n9>, p. I.

2. *Ibid.*, p. II.

3. *Ibid.*, p. III.

4. *Ibid.*

5. A. Lorry, « Les monographies des *Ouvriers européens* (1855, 1877-1879) et des *Ouvriers des deux mondes* (1857-1930). Inventaire et classification »..., p. 124.

des errata et une table des matières.

Le passage des volumes reliés aux fascicules trimestriels est acté par l'éditeur dans l'*Avertissement* liminaire du premier fascicule de la deuxième série :

« La nouvelle série des *Ouvriers des deux mondes* s'ouvre avec le présent fascicule. Le grand nombre des travaux soumis à la Société d'Économie sociale assure l'avenir et la régularité de la publication. *Les monographies paraîtront désormais en fascicules trimestriels*⁶. »

Dans l'*Avertissement* général de ce volume, édité en 1887 et placé immédiatement après la page de titre et le sommaire, le changement est à nouveau annoncé, mais également justifié :

« Quand la Société d'Économie sociale, en 1882, perdit celui qui avait réglé son rôle auprès de lui et après lui⁷, elle trouva le cinquième tome des *Ouvriers des deux mondes* arrêté au premier tiers de sa publication : elle acheva de le mettre au jour. Puis, voulant laisser bien distincts des travaux accomplis sous l'œil du maître, ceux dont il lui fallait dès lors prendre seule la responsabilité, *elle commença en juillet 1885, sous le même titre, une nouvelle série de monographies de familles paraissant par fascicules trimestriels*⁸. »

Deux versions de ce même volume sont à notre disposition. L'une résulte de la numérisation de l'exemplaire déposé à la Bibliothèque nationale de France à Paris et conservé par le département Philosophie, histoire, sciences de l'homme⁹, la seconde se trouve à la *John P. Robarts Research Library* de l'Université de Toronto et a été numérisée par *Internet Archive*¹⁰.

Des différences de composition existent. Ainsi, l'exemplaire de Toronto (fig. 1.1a, 1.1b, 1.1c) comporte deux feuillets qui ne sont pas présents dans celui de Paris (fig. 1.1d). Le premier porte au recto le titre des *Ouvriers des deux mondes* et au verso l'*Avertissement* que nous avons cité ci-dessus en premier, le second est la page de titre du fascicule avec un verso vierge. Dans l'exemplaire de la Bibliothèque nationale, la première page de la monographie n° 47 suit la fin de l'*Avertissement* général du volume.

Laisser aux propriétaires le soin de relier les fascicules pour composer le volume final revient à rompre l'unité codicologique de ce dernier, chaque nouvelle opération de reliure donnant naissance à un nouvel exemplaire. Les relieurs ont en effet pu choisir de

6. Alexis Delaire, « Paysan-paludier du Bourg de Batz (Loire-Inférieure - France) », dans *Les Ouvriers des deux mondes*. Paris, 1887 (Fascicule 1 (01)), p. 1-56, « Avertissement ». Nous soulignons.

7. Frédéric Le Play est mort le 5 avril 1882.

8. « Avertissement », dans *Les Ouvriers des deux mondes*. éd. Société internationale des études pratiques d'économie sociale, Paris, 1887, p. I-VIII, spec. p. II. Nous soulignons.

9. Consultable sur *Gallica* : <https://gallica.bnf.fr/ark:/12148/bpt6k54465138> (consulté le 20 août 2020).

10. Consultable à cette adresse : <https://archive.org/details/s2lesouvriersdes01sociuoft> (consulté le 20 août 2020).

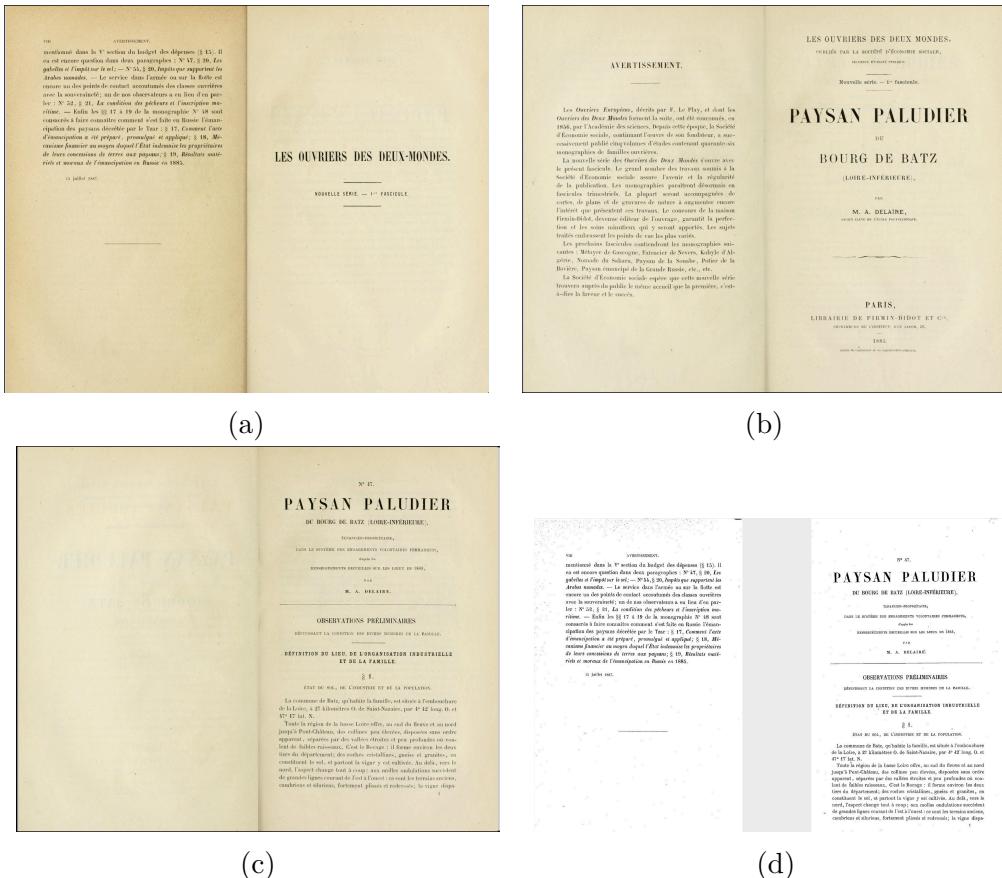


FIGURE 1.1 – Comparaison des exemplaires de Toronto et de Paris. (a) Toronto. Fin de l’*Avertissement* général du volume et début du fascicule. (b) Toronto. *Avertissement* du fascicule et page de titre de la monographie. (c) Toronto. Début de la monographie. (d) Paris. Fin de l’*Avertissement* général du volume et début de la monographie.

conserver ou de ne pas conserver les pages propres aux fascicules, à l’instar des feuillets de l’exemplaire de Toronto. Il n’existe donc pas un modèle ayant autorité dans la composition des *Ouvriers des deux mondes* — en dépit du fait que les volumes de la Bibliothèque nationale de France, issus du dépôt légal, ont dû être versés par l’éditeur, c’est-à-dire la Société internationale des études pratiques d’économie sociale.

Pour le projet *Time Us*, cela signifie que le corpus retenu pour le traitement informatisé et la publication finale ne sera qu’une version de l’œuvre que représente les *Ouvriers des deux mondes*.

1.2 La structure logique des monographies

Un élément remarquable traverse la centaine de monographies des *Ouvriers des deux mondes* : une même structure logique que les éditeurs ont tenté de maintenir tout au long de la publication.

Cette structure est l’incarnation de la méthodologie leplaysienne des monographies

de familles. Le terme *monographie*, « emprunté à l'histoire naturelle et à la médecine », qualifie au XIX^e siècle une « étude scientifique, minutieuse et détaillée, portant sur un objet ou un phénomène circonscrit »¹¹. La démarche monographique cherche à embrasser à travers l'observation directe des phénomènes inatteignables pour la démarche statistique, à laquelle elle s'oppose¹².

Avant les *Ouvriers des deux mondes*, Frédéric Le Play s'attelle aux *Ouvriers européens*, recueil paru en 1855 et réédité en de 1877 à 1879, contenant trente-six (1^{re} éd.) puis cinquante-sept (2^e éd.) monographies issues d'enquêtes menées entre 1833 et 1853¹³. Les *Ouvriers des deux mondes* se présentent comme une suite élargie à des espaces extra-européens de ce premier mouvement, dont ils reprennent la démarche, la structure et auquel ils font régulièrement référence par un système de renvoi. La structure logique des *Ouvriers des deux mondes*, dont le premier volume est publié en 1855, est donc déjà présente dans *Les Ouvriers européens*.

Elle procède d'une méthode d'observation que Le Play ne décrit qu'en 1862, au moment de la parution du quatrième volume des *Ouvriers des deux mondes*. Le titre de cette méthodologie précise néanmoins qu'elle a déjà fait ses preuves pour les *Ouvriers européens* : *Instruction sur la méthode d'observation dite des monographies de familles, propre à l'ouvrage intitulé Les ouvriers européens*¹⁴.

L'*Instruction* est divisée en quatre parties suivies d'une cinquième faisant office de conclusion. Dans la première, Le Play circonscrit son champ d'observation à la famille ouvrière¹⁵, érigée en « clef de compréhension des organisations sociales »¹⁶. La phase d'observation est ensuite divisée en étapes¹⁷, savoir l'observation des faits, l'interrogatoire de l'ouvrier et celui des personnes tiers¹⁸.

La troisième partie est celle où Le Play expose la structure logique qu'il utilise à

11. A. Savoye, « La monographie sociologique : jalons pour son histoire (1855-1974) », *Les Études sociales, Les monographies de famille de l'École de Le Play*–n° 131-132 (2000), p. 11-46, p. 12.

12. Voir en particulier la charge de F. Le Play contre les statisticiens dans son introduction aux *Ouvriers européens* de 1855 : « La méthode des statisticiens n'est pas l'observation directe des faits ; c'est la compilation et l'interprétation plus ou moins plausible de faits recueillis à des points de vue fort différents, étrangers pour la plupart à l'intérêt scientifique. Malgré leur généralité apparente et leur séduisante régularité, les statistiques ont médiocrement contribué au progrès de la science sociale » : F. Le Play, « Introduction. III, Appréciation des deux procédés communément employés pour observer les faits sociaux : supériorité des enquêtes directes sur les recherches scientifiques », dans *Les Ouvriers européens*, Paris, 1855, p. 9-12, URL : <https://gallica.bnf.fr/ark:/12148/bpt6k1057844n/f17>, p. 11.

13. Inventaire complet dans A. Lorry, « Les monographies des *Ouvriers européens* (1855, 1877-1879) et des *Ouvriers des deux mondes* (1857-1930). Inventaire et classification »..., p. 106-112.

14. F. Le Play, « Instruction sur la méthode d'observation dite des monographies de famille, propre à l'ouvrage intitulé *Les ouvriers européens* », dans *Les Ouvriers des deux mondes*, Paris, 1862 (série 1 (4)), p. 15-32, URL : <https://archive.org/details/lesouvriersdesde04sociuo0ft/page/14/>.

15. *Ibid.*, I., « Remarques préliminaires sur l'étude des faits sociaux et sur la méthode des monographies de familles », p. 15-16.

16. A. Savoye, « La monographie sociologique : jalons pour son histoire (1855-1974) »..., p. 15.

17. F. Le Play, « Instruction sur la méthode d'observation dite des monographies de famille, propre à l'ouvrage intitulé *Les ouvriers européens* »..., II., « Règles à suivre pour procéder à l'observation des faits sociaux », p. 17-19.

18. A. Savoye, « La monographie sociologique : jalons pour son histoire (1855-1974) »..., p. 16.

partir de 1855 et que ses continuateurs reprennent jusqu'en 1930. Elle est construite en miroir aux différentes phases de l'observation sur le terrain : à l'observation des faits répond la section des observations préliminaires, à l'interrogatoire de l'ouvrier celle des budgets et aux contacts avec les tiers celle des notes¹⁹. Le miroir est cependant légèrement déformant, afin de satisfaire au « hiatus » propre à la démarche leplaysienne, entre les « faits observés » et les « faits décris »²⁰. L'observation est en effet « contrainte par l'évidence des choses » là où la description, « parce qu'elle suppose de désigner précisément ce qu'on évoque, repose sur un calibrage sémantique »²¹. La structure logique des monographies de familles n'est donc pas qu'un simple plan de rédaction, elle procède d'une démarche intellectuelle réglée par Le Play et doit être appliquée pour garantir la validité de la démonstration.

Cette structure (Annexe A.2) possède trois niveaux de titre : les parties (A, B, C), les sections (I, II, etc.) et les paragraphes (§1, §2, etc.). Le dernier niveau est toujours atteint dans la partie B (*Observations préliminaires*), mais très rarement dans la partie C (*Notes*).

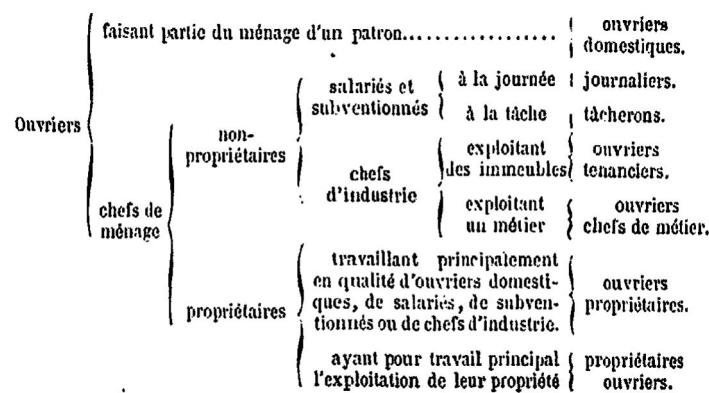


FIGURE 1.2 – « Tableau des sept situations principales que les ouvriers peuvent occuper successivement dans les quatre systèmes sociaux pour s'élever des rangs inférieurs de la hiérarchie industrielle à la condition de propriétaires ou de chefs d'industrie. », permettant de construire le titre d'une monographie de famille (capture d'écran de *Gallica*).

La première partie (A) est la page de titre. Le titre de la monographie est présenté comme le « résumé » de celle-ci, et doit pour cela contenir quatre éléments imprescriptibles : « 1° la profession de l'ouvrier, 2° la population dont il fait partie, 3° la nature des engagements qu'il contracte pour se procurer des moyens de travail, 4° la situation qu'il

19. F. Le Play, « Instruction sur la méthode d'observation dite des monographies de famille, propre à l'ouvrage intitulé *Les ouvriers européens* »..., III., « Précis des faits à observer — Établissement des budgets », p. 20-31.

20. S. Baciocchi et J. David, « IV. Ramifications épistémologiques. Décrire, définir, évaluer, comparer », *Les Études sociales, Frédéric Le Play : anthologie et correspondance*, n° 142-143-144 (II-2005-2006), p. 87-90, URL : https://www.academia.edu/5925885/Fr%C3%A9d%C3%A9ric_Le_Play_%C3%A9tude_sociale (visité le 11/08/2020), p. 87.

21. *Ibid.*, p. 87-88.

occupe dans l'organisation sociale caractérisée par cet engagement »²². Ainsi, en dépit du fait que Le Play ne considère pas explicitement la page de titre comme une composante de sa structure, l'importance qu'il donne au titre en lui-même a conduit le projet *Time Us* à l'interpréter comme telle. Un tableau pour construire ce titre est d'ailleurs mis à la disposition des monographies (fig. 1.2²³).

1.3 Continuité, discontinuité et erreurs d'impression dans la structure logique

En dépit de son aspect monolithique, la structure logique des monographies de familles subit plusieurs ajustements au cours de la publication. Certains, uniquement de circonstance, sont des évènements mineurs, là où d'autres sont des changements significatifs adoptés pour la suite de la publication.

La monographie n° 85 représente ainsi un tournant dans la numérotation des titres de paragraphe, sans pour autant changer leurs libellés (Annexe A.2). En effet, la numérotation cardinale est étendue au dernier paragraphe des budgets (*Comptes annexés aux budgets*) et remplace la numérotation alphabétique utilisée auparavant dans la partie *Notes*.

Pourquoi cette uniformisation intervient-elle à partir de cette monographie, en plein milieu du cinquième volume ? Nous avons vu plus haut que ce volume représentait un tournant dans la publication des *Ouvriers des deux mondes*, notamment en raison de la mort de Frédéric Le Play qui intervient avant qu'il ne soit achevé. Il est possible le départ du fondateur de la série ait permis d'opérer des changements dans la structure qu'il avait créée, ou bien que ce soit Le Play lui-même qui ait voulu ces changements. Quoi qu'il en soit, il y a une volonté claire de renforcer la solidité de la démonstration en rassemblant ses différents paragraphes en un seul bloc par le biais d'une numérotation cardinale continue. Le nombre moyen de paragraphes est de 21 ; la monographie la plus longue, composée de trente-cinq paragraphes, étant la n° 64, intitulée *Paysans corses en communauté, porchers-bergers des montagnes de Bastelica*²⁴.

22. F. Le Play, « Instruction sur la méthode d'observation dite des monographies de famille, propre à l'ouvrage intitulé *Les ouvriers européens* »..., p. 20.

23. *Ibid.*, p. 21 .

24. Maximilien Bigot et François Escard, « Paysans corses en communauté, porchers-bergers des montagnes de Bastelica », dans *Les Ouvriers des deux mondes*, Paris, 1890 (série 2e (2)), p. 433-524.

Chapitre 2

Des numérisations multiples

Début du deuxième chapitre

Chapitre 3

Un encodage automatique

Début du troisième chapitre

Annexes

A. *Les Ouvriers des deux mondes*

A.1 Liste des monographies et fichiers correspondant

La date de publication du volume est indiquée entre parenthèses, tandis que les dates de parution des fascicules le sont entre crochets droits.

A.1.1 Série 1, vol. 1 (1857).

Id	Intitulé	§	Fichier
401a	Page de titre	–	s1t1_chapt_1.xml
402a	Avertissement. Considérations générales sur la Société internationale des études pratiques d'économie sociale. Son but et ses moyens d'action.	–	s1t1_chapt_2.xml
403a	Institution. Société internationale des études pratiques d'économie sociale. Fondation et premiers travaux.	–	s1t1_chapt_3.xml
404a	Définitions par ordre alphabétiques des termes à employer dans les monographies, pour désigner les ouvriers, leurs moyens d'existence, et les rapports qui les unissent soit entre eux, soit avec les autres classes.	–	s1t1_chapt_4.xml
405a	Explications des signes de renvoi et des abréviations.	–	s1t1_chapt_5.xml
001a	Charpentier de Paris (Seine - France), de la Corporation des compagnons du Devoir	–	s1t1_chapt_6.xml
002a	Manœuvre-Agriculteur de la Champagne pouilleuse (Marne - France)	–	s1t1_chapt_7.xml
003a	Paysans en communauté du Lavedan (Hautes-Pyrénées - France)	–	s1t1_chapt_8.xml
004a	Paysan du Labourd (Basses-Pyrénées - France)	–	s1t1_chapt_9.xml

005a	Métayer de la banlieue de Florence (Grand-Duché de Toscane)	–	s1t1_chapt_10.xml
006a	Nourrisseur de vaches de la banlieue de Londres (Middlesex - Angleterre)	–	s1t1_chapt_11.xml
007a	Tisseur en châles de la fabrique urbaine collective de Paris (Seine - France)	–	s1t1_chapt_12.xml
008a	Manœuvre-agriculteur du comté de Nottingham (Angleterre)	–	s1t1_chapt_13.xml
009a	Pêcheur côtier, maître de barque de Saint-Sébastien (Guipuscoa - Espagne)	–	s1t1_chapt_14.xml
406a	Tables alphabétique et analytique des matières contenues dans ce tome premier.	–	s1t1_chapt_15.xml
407a	Liste des monographies destinées aux prochaines publications de la société d'économie sociale.	–	s1t1_chapt_16.xml
408a	Tables des matières contenues dans ce tome premier	–	s1t1_chapt_17.xml

A.1.2 Série 1, vol. 2 (1858).

Id	Intitulé	§	Fichier
409a	Page de titre	–	s1t2_chapt_1.xml
410a	Avertissement	–	s1t2_chapt_2.xml
010a	Ferblantier, couvreur et vitrier d'Aix-les-Bains (Savoie - États Sardes)	–	s1t2_chapt_3.xml
011a	Carrier des environs de Paris (Seine - France)	–	s1t2_chapt_4.xml
012a	Menuisier-charpentier (Nedjar) de Tanger (Province de Tanger - Maroc)	–	s1t2_chapt_5.xml
013a	Tailleur d'habits de Paris (Seine - France)	–	s1t2_chapt_6.xml
014a	Compositeur-typographe de Bruxelles (Brabant - Belgique)	–	s1t2_chapt_7.xml
015a	Décapeur d'outils en acier de la fabrique d'Hérimoncourt (Doubs - France)	–	s1t2_chapt_8.xml
016a	Monteur d'outils en acier de la fabrique d'Hérimoncourt (Doubs - France)	–	s1t2_chapt_9.xml
017a	Porteur d'eau de Paris (Seine - France)	–	s1t2_chapt_10.xml

018a	Paysans en communauté et en polygamie de Bousrah (Esky Cham), dans le pays de Haouran (Syrie - Empire Ottoman)	-	s1t2_chapt_11.xml
019a	Débardeur et piocheur de craie de la banlieue de Paris (Seine - France)	-	s1t2_chapt_12.xml
411a	Tables alphabétique et analytique des matières contenues dans ce tome second	-	s1t2_chapt_13.xml
412a	Errata	-	s1t2_chapt_14.xml
413a	Tables des matières contenues dans ce tome second	-	s1t2_chapt_15.xml

A.1.3 Série 1, vol. 3 (1861).

Id	Intitulé	§	Fichier
414a	Page de titre	-	s1t3_chapt_1.xml
415a	Avertissement	-	s1t3_chapt_2.xml
416a	Rapport. Société d'économie sociale. Travaux de 1859-1860	-	s1t3_chapt_3.xml
417a	Liste générale des membres de la Société internationale des études pratiques d'économie sociale	-	s1t3_chapt_4.xml
020a	Brodeuses des Vosges (Vosges - France)	-	s1t3_chapt_5.xml
021a	Paysan et savonnier de la Basse-Provence (Bouches-du-Rhône - France)	-	s1t3_chapt_6.xml
022a	Mineur des Placers du comté de Mariposa (Californie - États-Unis)	-	s1t3_chapt_7.xml
023a	Manœuvre-vigneron de l'Aunis (Charente-inférieure - France)	-	s1t3_chapt_8.xml
024a	Lingère de Lille (Nord - France)	-	s1t3_chapt_9.xml
025a	Parfumeur de Tunis (Régence de Tunis - Afrique) du bazar appelé : El Attharin-el-kebar (les grands parfumeurs)	-	s1t3_chapt_10.xml
026a	Instituteur primaire d'une commune rurale de la Normandie (Eure - France)	-	s1t3_chapt_11.xml
027a	Manœuvre à famille nombreuse de Paris (Seine - France)	-	s1t3_chapt_12.xml
028a	Fondeur de plomb des Alpes Apuanes (Toscane - Italie)	-	s1t3_chapt_13.xml

418a	Tables alphabétique et analytique des matières contenues dans ce tome troisième	—	s1t3_chapt_14.xml
419a	Errata de ce tome troisième	—	s1t3_chapt_15.xml
420a	Tables des matières contenues dans ce tome troisième	—	s1t3_chapt_16.xml

A.1.4 Série 1, vol. 4 (1862).

Id	Intitulé	§	Fichier
421a	Page de titre	—	s1t4_chapt_1.xml
422a	Explications des signes de renvoi et des abréviations.	—	s1t4_chapt_2.xml
423a	Avertissement	—	s1t4_chapt_3.xml
424a	Rapport. Société d'économie sociale. Travaux de 1860-1861	—	s1t4_chapt_4.xml
425a	Instruction. Méthode d'observation des monographies de famille propre à l'ouvrage intitulé Les ouvriers européens	—	s1t4_chapt_5.xml
426a	Histoire de la famille. Prix fondé par M. le baron de Damas. Par la société d'économie sociale	—	s1t4_chapt_6.xml
029a	Paysan d'un village à banlieue morcelée du Laonnais (Aisne - France)	—	s1t4_chapt_7.xml
030a	Paysans en communauté du Ning-Po-Fou (province de Tché-Kian - Chine)	—	s1t4_chapt_8.xml
031a	Mulâtre affranchi de l'Ile de la Réunion (Océan Indien)	—	s1t4_chapt_9.xml
032a	Manœuvre-vigneron de la Basse-Bourgogne (Yonne - France)	—	s1t4_chapt_10.xml
033a	Compositeur-typographe de Paris (Seine - France)	—	s1t4_chapt_11.xml
034a	Auvergnat brocanteur en boutique à Paris (Seine - France)	—	s1t4_chapt_12.xml
035a	Mineur de la Maremme de Toscane (Toscane - Italie)	—	s1t4_chapt_13.xml
036a	Tisserand des Vosges (Haut-Rhin - France)	—	s1t4_chapt_14.xml
037a	Pêcheur côtier, maître de barques, de Marken (Hollande septentrionale - Pays-Bas)	—	s1t4_chapt_15.xml

427a	Société internationale des études pratiques d'économie sociale. Officiers composants les conseils d'administration et de surveillance pour la session 1863-1864.	-	s1t4_chapt_16.xml
428a	Liste générale des membres de la Société internationale des études pratiques d'économie sociale au 1er août 1863	-	s1t4_chapt_17.xml
429a	Tables alphabétique et analytique des matières contenues dans ce tome quatrième	-	s1t4_chapt_18.xml
430a	Errata de ce tome quatrième	-	s1t4_chapt_19.xml
431a	Tables des matières contenues dans ce tome quatrième	-	s1t4_chapt_20.xml

A.1.5 Série 1, vol. 5 [1875, 1883, 1884] (1885).

Id	Intitulé	§	Fichier
432a	Page de titre	-	s1t5_chapt_1.xml
433a	Oeuvres de F. Le Play	-	s1t5_chapt_2.xml
434a	Sommaire. Monographies de familles publiées dans ce volume	-	s1t5_chapt_3.xml
435a	Avertissement	-	s1t5_chapt_4.xml
436a	Explications des signes de renvoi et des abréviations employés dans le cours de cet ouvrage	-	s1t5_chapt_5.xml
038a	Fermiers à communauté taisable du Nivernais (Saône-et-Loire - France)	-	s1t5_chapt_6.xml
039a	Paysan de Saint-Irénée (Bas-Canada - Amérique du Nord)	-	s1t5_chapt_7.xml
040a	L'Ouvrier éventailiste de Sainte-Geneviève (Oise - France)	-	s1t5_chapt_8.xml
041a	Ouvrier cordonnier de Malakoff (Seine - France)	-	s1t5_chapt_9.xml
041b	Précis d'une monographie ayant pour objet un chiffonnier instable et, par alternance, mégissier fumiste et brossier de Paris (France - Seine)	-	<i>In supra.</i>
042a	Serrurier-forgeron de Paris (Seine - France)	-	s1t5_chapt_10.xml
042b	Précis d'une monographie ayant pour objet le monteur en bronze de Paris	-	<i>In supra.</i>

043a	Brigadier de la Garde républicaine de Paris (Seine - France)	-	s1t5_chapt_11.xml
044a	Paysan-résinier de Lévignacq (Landes - France)	-	s1t5_chapt_12.xml
045a	Bûcheron usager de l'ancien Comté de Dabo (Lorraine allemande)	-	s1t5_chapt_13.xml
046a	Paysans en communauté et colporteurs émigrants de Tabou-Douchd-El-Baar (Grande Kabylie - Province d'Alger)	-	s1t5_chapt_14.xml
437a	Société d'économie sociale. Conseil d'administration pour l'année 1885	-	s1t5_chapt_15.xml
438a	Liste générale des membres de la Société d'économie sociale au 15 mars 1885	-	s1t5_chapt_16.xml
439a	Tables alphabétique et analytique des matières contenues dans ce tome cinquième	-	s1t5_chapt_17.xml
440a	Tables des matières contenues dans ce tome cinquième	-	s1t5_chapt_18.xml

A.1.6 Série 2, vol. 1 [1885-1887] (1887).

Id	Intitulé	§	Fichier
441a	Page de titre	-	s2t1_chapt_1.xml
442a	Sommaire des monographies de familles publiées dans ce volume	-	s2t1_chapt_2.xml
443a	Avertissement sur ce premier volume, deuxième série des Ouvriers des deux mondes	-	s2t1_chapt_3.xml
047a	Paysan-paludier du Bourg de Batz (Loire-Inférieure - France)	-	s2t1_chapt_4.xml
048b	Précis d'une monographie de l'armurier des manufactures impériales de Toula (Grande-Russie)	-	s2t1_chapt_6.xml
048a	Bordiers émancipés en communauté rurale de la Grande-Russie	-	s2t1_chapt_5.xml
049a	Charron des forges et fonderies de Montataire (Oise - France)	-	s2t1_chapt_16.xml
050a	Faienciers de Nevers (Nièvre - France)	-	s2t1_chapt_17.xml
051a	Cultivateur-maraîcher de Deuil (Seine-et-Oise - France)	-	s2t1_chapt_18.xml

052a	Pêcheur-côtier, maître de barque, de Martigues (Bouches-du-Rhône - France)	-	s2t1_chapt_19.xml
053a	Métayer à famille-souche du pays d'Horte (Landes - France)	-	s2t1_chapt_20.xml
054a	Arabes pasteurs nomades de la tribu des Larbas (Région saharienne de l'Algérie)	-	s2t1_chapt_21.xml
055a	Gantier de Grenoble (Isère - France)	-	s2t1_chapt_22.xml
444a	Tables alphabétique et analytique des matières contenues dans le présent volume	-	s2t1_chapt_24.xml
445a	Table des matières dans ce tome premier (deuxième série)	-	s2t1_chapt_25.xml

A.1.7 Série 2, vol. 2 [1887-1889] (1890).

Id	Intitulé	§	Fichier
446a	Page de titre	-	s2t2_chapt_1.xml
447a	Page de titre	-	s2t2_chapt_2.xml
448a	Sommaire des monographies de familles publiées dans ce volume	-	s2t2_chapt_3.xml
449a	Avertissement sur ce deuxième tome de la deuxième série des Ouvriers des deux mondes	-	s2t2_chapt_4.xml
056a	Tourneur-mécanicien des usines de la Société Cockerill, de Seraing (Belgique)	-	s2t2_chapt_5.xml
057a	Bordier (Fellah) berbère de la Grande-Kabylie (Province d'Alger)	-	s2t2_chapt_6.xml
057b	Précis d'une monographie du paysan colon du Sahel (Algérie)	-	s2t2_chapt_7.xml
058a	Pêcheur côtier d'Heyst (Flandre occidentale - Belgique)	-	s2t2_chapt_8.xml
058b	Précis d'une monographie du pêcheur côtier, maître de barque, d'Étretat (Seine-Inférieure - France)	-	s2t2_chapt_9.xml
059a	Paysan-métayer de la Basse Provence (Bouches-du-Rhône - France)	-	s2t2_chapt_10.xml
059b	Précis d'une monographie du paysan et maçon émigrant de la Marche (Creuse - France)	-	s2t2_chapt_11.xml

060a	Mineur silésien du bassin houiller de la Ruhr (Prusse rhénane - Allemagne)	-	s2t2_chapt_12.xml
061a	Mineur des soufrières de Lercara (Province de Palerme - Sicile)	-	s2t2_chapt_13.xml
062a	Tailleur de Silex et vigneron de l'Orléanais (Loir-et-Cher - France)	-	s2t2_chapt_14.xml
063a	Vigneron précariste et métayer de Valmontone (Province de Rome - Italie)	-	s2t2_chapt_15.xml
064a	Paysans corses en communauté, porchers-bergers des montagnes de Bastelica	-	s2t2_chapt_16.xml
450a	Tables alphabétique et analytique des matières contenues dans le présent tome, avec index explicatif des mots employés dans un sens propre à l'économie sociale	-	s2t2_chapt_17-1.xml
450b	Table des matières dans ce tome deuxième (deuxième série)	-	s2t2_chapt_17-2.xml

A.1.8 Série 2, vol. 3 [1890-1892] (1892).

Id	Intitulé	§	Fichier
451a	Page de titre	-	s2t3_chapt_1.xml
452a	Page de titre	-	s2t3_chapt_2.xml
453a	Sommaire des monographies de familles publiées dans ce volume	-	s2t3_chapt_3.xml
454a	Avertissement sur ce troisième tome de la deuxième série des Ouvriers des deux mondes	-	s2t3_chapt_4.xml
065a	Métayers en communauté du Confolentais (Charente - France)	-	s2t3_chapt_5.xml
066a	Vignerons de Ribeaupillé (Alsace)	-	s2t3_chapt_6.xml
066b	Précis d'une monographie du pêcheur-côtier du Finmark (Laponie - Norvège)	-	s2t3_chapt_7.xml
066c	Précis d'une monographie d'un tisserand d'Hilversum (Hollande septentrionale - Pays-Bas)	-	s2t3_chapt_14.xml
067a	Tisserand de la fabrique collective de Gand (Flandre orientale - Belgique)	-	s2t3_chapt_28.xml

068a	Paysan agriculteur de Torremaggiore (Province de Foggia - Italie)	-	s2t3_chapt_29.xml
069a	Tanneur de Nottingham (Angleterre)	-	s2t3_chapt_30.xml
070a	Charpentier indépendant de Paris (Seine - France)	-	s2t3_chapt_31.xml
071a	Conducteur-typographe de l'agglomération bruxelloise (Brabant - Belgique)	-	s2t3_chapt_32.xml
072a	Coutelier de la fabrique collective de Gembloux (Province de Namur - Belgique)	-	s2t3_chapt_33.xml
455a	Tables alphabétique et analytique des matières contenues dans le présent tome, avec index explicatif des mots employés dans un sens propre à l'économie sociale	-	s2t3_chapt_34.xml
456a	Table des matières dans ce tome troisième (deuxième série)	-	s2t3_chapt_35.xml

A.1.9 Série 2, vol. 4 [1892-1895] (1895).

Id	Intitulé	§	Fichier
457a	Page de titre	-	s2t4_chapt_1.xml
458a	Page de titre	-	s2t4_chapt_2.xml
459a	Sommaire des monographies de familles publiées dans ce volume	-	s2t4_chapt_3.xml
460a	Avertissement sur ce quatrième volume de la deuxième série	-	s2t4_chapt_4.xml
073a	Ajusteur-surveillant de l'usine de Guise (Aisne - France)	-	s2t4_chapt_5.xml
074a	Ébéniste parisien de haut luxe (Seine - France)	-	s2t4_chapt_6.xml
075a	Métayer de l'Ouest du Texas (États-Unis d'Amérique)	-	s2t4_chapt_7.xml
076a	Ouvrière mouleuse en cartonnage d'une fabrique collective de jouets parisiens (Seine - France)	-	s2t4_chapt_8.xml
077a	Savetier de Bâle (Suisse)	-	s2t4_chapt_9.xml
078a	Ouvrier-employé de la fabrique coopérative de papiers d'Angoulême (Charente - France)	-	s2t4_chapt_10.xml
079a	Tisseur de San Leucio (Province de Caserte - Italie)	-	s2t4_chapt_11.xml

080a	Fermiers montagnards du Haut-Forez (Loire - France)	-	s2t4_chapt_12.xml
081a	Allumeur de réverbères de Nancy (Meurthe-et-Moselle - France)	-	s2t4_chapt_13.xml
461a	Tables alphabétique et analytique des matières contenues dans le présent tome, avec index explicatif des mots employés dans un sens propre à l'économie sociale	-	s2t4_chapt_14.xml
462a	Table des matières contenues dans ce tome quatrième (deuxième série)	-	s2t4_chapt_15.xml

A.1.10 Série 2, vol. 5 [1895-1899] (1899).

Id	Intitulé	§	Fichier
463a	Page de titre	-	s2t5_chapt_1.xml
464a	Société d'économie sociale [nota : liste des publications]	-	s2t5_chapt_2.xml
465a	Sommaire des monographies de familles publiées dans ce volume	-	s2t5_chapt_3.xml
466a	Avertissement sur ce cinquième tome de la deuxième série	-	s2t5_chapt_4.xml
082a	Ouvrier garnisseur de canons de fusils de la fabrique collective d'armes à feu de Liège (Liège - Belgique)	-	s2t5_chapt_5.xml
083a	Fileur en peigné et réglleur de métier de la Manufacture du Val-des-Bois (Marne - France)	-	s2t5_chapt_6.xml
084a	Cordonnier d'Iseghem (Flandre Occidentale - Belgique)	-	s2t5_chapt_7.xml
085a	Paysan métayer (Contadino mezzajuolo) de Roccasancasciano (Romagne Toscane - Italie)	-	s2t5_chapt_8.xml
085b	Précis d'une monographie d'un ouvrier agriculteur de la campagne de Ravenne (Romagne - Italie)	-	s2t5_chapt_9.xml
086a	Mineur des mines de houille du Pas-de-Calais (France)	-	s2t5_chapt_10.xml
087a	Agriculteur du Pas-de-Calais (France)	-	s2t5_chapt_11.xml

088a	Serrurier-forgeron du quartier de Picpus, à Paris (France)	-	s2t5_chapt_12.xml
088b	Précis d'une monographie du serrurier poseur de persiennes en fer de Paris	-	s2t5_chapt_13.xml
089a	Piqueur sociétaire de la Mine aux Mineurs de Monthieux (Loire - France)	-	s2t5_chapt_14.xml
090a	Petit fonctionnaire de Pnom-Penh (Cambodge)	-	s2t5_chapt_15.xml
090b	Précis d'une monographie d'un manœuvre-coolie de Pnom-Penh (Cambodge)	-	s2t5_chapt_16.xml
091a	Métayer de Corrèze (Bas Limousin - France)	-	s2t5_chapt_17.xml
467a	Tables alphabétique et analytique des matières contenues dans le présent tome, avec index explicatif des mots employés dans un sens propre à l'économie sociale	-	s2t5_chapt_18-1.xml
467b	Table des matières dans ce tome cinquième	-	s2t5_chapt_18-2.xml

A.1.11 Série 3, vol. 1 [1900-1904] (1904).

Id	Intitulé	§	Fichier
468a	[Fichier sans texte]	-	s3t1_chapt_1.xml
472a	La société générale des papeteries du Limousins	-	s3t1_chapt_2.xml
092a	Fermier normand de Jersey	-	s3t1_chapt_3.xml
092b	Précis d'une monographie d'un pêcheur-côtier, maître de barques, de l'archipel Chusan (Chine)	-	s3t1_chapt_4.xml
093a	Aveugle accordeur de pianos de Levallois-Perret (Seine - France)	-	s3t1_chapt_5.xml
094a	Bouilleur de cru du Bas-Pays de Cognac (Charente - France)	-	s3t1_chapt_6.xml
095a	Mineur du bassin houiller du Couchant de Mons (Borinage - Belgique)	-	s3t1_chapt_7.xml
096a	Fellah de Karnak (Haute-Egypte)	-	s3t1_chapt_8.xml
097a	Tisserand d'usine de Gladbach (Prusse rhénane)	-	s3t1_chapt_9.xml
098a	Décoreuse de porcelaine de Limoges (Haute-Vienne - France)	-	s3t1_chapt_10.xml
099a	Cantonnier-poseur de voie du chemin de fer du Nord à Paris	-	s3t1_chapt_11.xml

A.1.12 Série 3, vol. 2 [1904-1908] (1908).

Id	Intitulé	§	Fichier
469a	Page de titre	—	s3t2_chapt_1.xml
100a	Cordonnier de la fabrique collective de Binche (Province de Hainaut - Belgique)	—	s3t2_chapt_2.xml
101a	Compositeur typographe de Québec (Canada - Amérique du Nord)	—	s3t2_chapt_3.xml
102a	Ardoisier du bassin d'Herbeumont (Belgique)	—	s3t2_chapt_4.xml
103a	Commis à l'administration centrale des chemins de fer de l'État belge (Schaerbeek-Bruxelles - Belgique)	—	s3t2_chapt_5.xml
104a	Teinturier de ganterie et gantières de Saint-Junien (Haute-Vienne - France)	—	s3t2_chapt_6.xml
105a	Jardinier-plantier de Gasseras (Commune de Montauban, Tarn-et-Garonne - France)	—	s3t2_chapt_7.xml
106a	Corsetière du Raincy (banlieue de Paris - France)	—	s3t2_chapt_8.xml
107a	Étameur sur fer-blanc des usines de Commentry (Allier - France)	—	s3t2_chapt_9.xml
473a	Usine hydraulique d'éclairage et de transport de force	—	s3t2_chapt_10.xml

A.1.13 Série 3, vol. 3 [1908-1913] (1913).

Id	Intitulé	§	Fichier
470a	Page de titre	—	s3t3_chapt_1.xml
108a	Paysan cultivateur du Ruvo di Puglia (Province de Bari - Italie, 1903)	—	s3t3_chapt_2.xml

A.1.14 Série 3, vol. 3bis [1928-1930] (1930).

Id	Intitulé	§	Fichier
471a	Page de titre	—	s3t3-bis_chapt_1.xml

A.2 Structure logique

- A. *Titre.*
- B. *Observations préliminaires définissant la condition des divers membres de la famille.*
 - I. *Définition du lieu, de l'organisation industrielle et de la famille.*
 - § 1. *État du sol, de l'industrie et de la population.*
 - § 2. *État civil de la famille.*
 - § 3. *Religion et habitudes morales.*
 - § 4. *Hygiène et services de santé.*
 - § 5. *Rang de la famille.*
 - II. *Moyens d'existence de la famille.*
 - § 6. *Propriétés.*
 - § 7. *Subventions.*
 - § 8. *Travaux et industries.*
 - III. *Mode d'existence de la famille.*
 - § 9. *Aliments et repas.*
 - § 10. *Habitation, mobilier et vêtements.*
 - § 11. *Récréations.*
 - IV. *Histoire de la famille.*
 - § 12. *Phases principales de l'existence.*
 - § 13. *Mœurs et institutions assurant le bien-être physique et moral de la famille.*
 - V. (*Budget domestique annuel*¹).
 - § 14. *Budget des recettes de l'année.*
 - § 15. *Budget des dépenses de l'année.*

Comptes annexés aux budgets (n° 1 à 84) puis § 16. Comptes annexés aux budgets.
- C. *Notes* (n° 1 à 84) puis *Éléments divers de la constitution sociale.*
 - (A) (*titre du paragraphe*) (n° 1 à 84) puis § 17. (*titre du paragraphe*).
 - (B) (*titre du paragraphe*) (n° 1 à 84) puis § 18. (*titre du paragraphe*).
 - etc.*

1. Cette section ne possède un titre que dans huit monographies

Table des matières

Résumé	iii
Liste des sigles et abréviations	v
Préambule : un stage confiné ?	vii
Introduction	xi
Bibliographie	xix
Bibliographie générale	xx
<i>Les Ouvriers des deux mondes</i>	xxi
I Un corpus déjà structuré	1
1 Un corpus d'imprimés	5
1.1 Une publication sur une soixantaine d'années	5
1.2 La structure logique des monographies	7
1.3 Continuité, discontinuité et erreurs d'impression	10
2 Des numérisations multiples	11
3 Un encodage automatique	13
Annexes	17
A. <i>Les Ouvriers des deux mondes</i>	17
A.1 Liste des monographies et fichiers correspondant	17
A.1.1 Série 1, vol. 1 (1857).	17
A.1.2 Série 1, vol. 2 (1858).	18
A.1.3 Série 1, vol. 3 (1861).	19
A.1.4 Série 1, vol. 4 (1862).	20
A.1.5 Série 1, vol. 5 [1875, 1883, 1884] (1885).	21

A.1.6 Série 2, vol. 1 [1885-1887] (1887).	22
A.1.7 Série 2, vol. 2 [1887-1889] (1890).	23
A.1.8 Série 2, vol. 3 [1890-1892] (1892).	24
A.1.9 Série 2, vol. 4 [1892-1895] (1895).	25
A.1.10 Série 2, vol. 5 [1895-1899] (1899).	26
A.1.11 Série 3, vol. 1 [1900-1904] (1904).	27
A.1.12 Série 3, vol. 2 [1904-1908] (1908).	28
A.1.13 Série 3, vol. 3 [1908-1913] (1913).	28
A.1.14 Série 3, vol. 3bis [1928-1930] (1930).	28
A.2 Structure logique	29
Table des matières	31