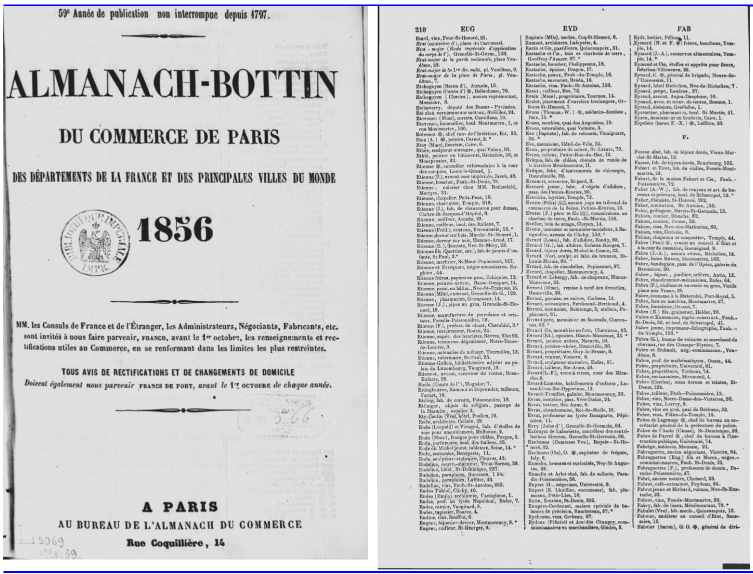
Repopulating Paris: massive extraction of 4 Million addresses from city directories between 1839 and 1922.

Isabella di Lenardo (isabella.dilenardo@epfl.ch), DHLAB-EPFL, Switzerland and Raphaël Barman (raphael.barman@epfl.ch), Institut National d'Histoire de l'Art and Albane Descombes (albane.descombes@epfl.ch), Institut National d'Histoire de l'Art and Frédéric Kaplan (frederic.kaplan@epfl.ch), DHLAB-EPFL, Switzerland

1. Introduction

In 1839, in Paris, the Maison Didot bought the Bottin company. Sébastien Bottin trained as a statistician was the initiator of a high impact yearly publication, called "Almanachs" containing the listing of residents, businesses and institutions, arranged geographically, alphabetically and by activity typologies (Fig. 1). These regular publications encountered a great success. In 1820, the Parisian Bottin Almanach contained more than 50 000 addresses and until the end of the 20th century the word "Bottin" was the colloquial term to designate a city directory in France. The publication of the "Didot-Bottin" continued at an annual rhythm, mapping the evolution of the active population of Paris and other cities in France.



59e Année de publication non interrompue depuis 1797.

ALMANACH-BOTTIN

DU COMMERCE DE PARIS

DES DÉPARTEMENTS DE LA FRANCE ET DES PRINCIPALES VILLES DU MONDE

BIBLIOTHÈQUE IMPÉRIALE IMPR.

1856

MM. les Consuls de France et de l'Étranger, les Administrateurs, Négociants, Fabricants, etc. sont invités à nous faire parvenir, FRANCO, avant le 1er octobre, les renseignements et rectifications utiles au Commerce, en se renfermant dans les limites les plus restreintes.

TOUS AVIS DE RECTIFICATIONS ET DE CHANGEMENTS DE DOMICILE
Doivent également nous parvenir FRANCS DE PORT, *avant le* 1er OCTOBRE *de chaque année.*

A PARIS
AU BUREAU DE L'ALMANACH DU COMMERCE
Rue Coquillière, 14

210 EUG EYD FAB

Etard, vins, Four-St-Honoré, 21.
Etat (ministère d'), place du Carrousel.
Etat-major (*Ecole impériale d'application du corps de l'*), Grenelle-St-Germ., 138.
Etat-major de la garde nationale, place Vendôme, 22.
État-major de la 1re *div. milit.* pl. Vendôme, 9.
Etat-major de la place de Paris, pl. Vendôme, 7.
Etchegoyen (Baron d'), Aumale, 15.
Etchegoyen (Comte d') ✻, Bellechasse, 70.
Etchegoyen (Charles), ancien représentant, Monsieur, 8.
Etcheverry, député des Basses-Pyrénées.
Été aîné, vernisseur sur métaux, St-Gilles, 24.
Etevenon (Mme), corsets, Castellane, 10.
Etevenon, limonadier, boul. Montmartre, 1, et rue Montmartre, 160.
Etévenot ✻, chef retr. de l'Intérieur, Est, 35.
Etex (A.) ✻, peintre, Carnot, 2. *
Etey (Mme), fleuriste, Caire, 6.
Ethée, sculpteur statuaire, quai Valmy, 83.
Ethié, peintre en bâtiments, Richelieu, 26, et Montpensier, 23.
Etienne ✻, conseiller référendaire à la cour des comptes, Louis-le-Grand, 1.
Etienne (P.), avocat cour impériale, Jacob, 48.
Etienne, boucher, Faub.-St-Denis, 79.
Etienne, caissier chez MM. Rothschild, Martyrs, 31.
Etienne, chapelier. Petit-Pont, 18.
Etienne, charcutier, Temple, 219.
Etienne (J.), fab. de chaussures pour dames, Cloître-St-Jacques-l'Hôpital, 8.
Etienne, coiffeur, Arcade, 29.
Etienne, coiffeur, boul. des Italiens, 7.
Etienne (Fréd.), cristaux, Ferronnerie, 13. *
Etienne, doreur sur bois, Marché-St-Honoré, 1.
Etienne, doreur sur bois, Homme-Armé, 17.
Etienne (E.), fleuriste, Nve-St-Méry, 23.
Etienne fils (Quettier, suc.), fab. de jouets d'enfants, St-Paul, 2.*
Etienne, marbrier, St-Maur-Popincourt, 157.
Etienne et Errequeta, négts commissres. Enghien, 44.
Etienne frères, papiers en gros, Echiquier, 12.
Etienne, peintre-artiste, Basse-Rempart, 14.
Etienne, peint. en bâtim., Nve-St-François, 16.
Etienne (Mlle), externat, Grenelle-St-G., 122.
Etienne, pharmacien, Grammont, 14.
Etienne (J.), pipes en gros, Grenelle-St-Honoré, 19.
Etienne, manufacture de porcelaine et cristaux, Paradis-Poissonnière, 29.
Etienne (F.), profess. de chant, Cherubini, 3.*
Etienne, restaurateur, Bouloi, 24.
Etienne, supér. des lazaristes, Sèvres, 93 et 95.
Etienne, teinturier-dégraisseur, Notre-Dame-de-Lorette, 9.
Etienne, ustensiles de ménage, Tournelles, 13.
Etienne, vétérinaire, St-Paul, 25.
Etienne-Gallois, bibliothécaire adjoint au palais du Luxembourg, Vaugirard, 19.
Etiennot, avocat, receveur de rentes, Bons-Enfants, 26.
Etoile (Comte de l'), Mogador, 7.
Ettinghausen, Ramond et Deprendez, tailleurs, Favart, 18.
Ettling, fab. de corsets, Poissonnière, 18.
Ettinger, objets de religion, passage de la Marmite, escalier A.
Ety-Coutin (Vve), hôtel, Poulies, 10.
Eude, architecte, Colisée, 29.
Eude (Léopold) et Vieugué, fab. d'étoffes de soie pour ameublement, Mulhouse, 2.
Eude (Mme), franges pour châles, Forges, 5.
Eude, parfumerie, boul. des Italiens, 35.
Eude dit Michel jeune, tableaux, Seine, 14. *
Eude, costumier, Bonaparte, 11.
Eude, sculpteur-statuaire, Fleurus, 42.
Eudeline, couvr.-zingueur, Trois-Bornes, 39.
Eudeline, hôtel, St-Dominique, 237.
Eudeline, parapluies, Suresnes, 1 *bis*.
Eudeline, parapluies, Laffitte, 43.
Eudeline, vins, Faub.-St-Antoine, 225.
Eudes (l'abbé), Clichy, 46.
Eudes (Emile) architecte, Castiglione, 1.
Eudes, prof. au lycée Napoléon, Enfer, 7.
Eudes, rentier, Vaugirard, 9.
Eudes, tapissier, Bourse, 4.
Eudine, vins, Soufflot, 6.
Eugène, bijoutier-doreur, Montmorency, 9. *
Eugène, coiffeur, St-Georges, 9.
Eugénie (Mlle), modes, Coq-St-Honoré, 6.
Eumont, architecte, Lafayette, 4.
Eurin et Cie, pastilleurs, Quincampoix, 31.
Eustache et Cie, bois et charbons de terre, Geoffroy-l'Asnier, 27. *
Eustache, boucher, Phélippeaux, 18.
Eustache, épicier, Dragon, 17.
Eustache, peaux, Faub.-du-Temple, 16.
Eustache, serrurier, Bréda, 18.
Eustache, vins, Faub.-St-Antoine, 158.
Euzet, coiffeur, Bac, 73.
Evain (Mme), propriétaire, Tournon, 14.
Evalet, placement d'ouvriers boulangers, Orléans-St-Honoré, 7.
Evans (Thomas.-W.) ✻, médecin-dentiste, Paix, 15. *
Evans, meubles, quai des Augustins, 19.
Evans, naturaliste, quai Voltaire, 3.
Eve (Baptiste), fab. de robinets, Vinaigriers, 36. *
Eve, menuisier, Hôtel-de-Ville, 56.
Even, propriétaire de mines, St-Lazare, 79.
Eveno, relieur, Petite-Rue-du-Bac, 12.
Evêque, fab. de châles, chemin de ronde de la barrière Ménilmontant, 11.
Evêque, fabr. d'instruments de chirurgie, Hautefeuille, 32.
Everaert, serrurier, Regard, 2.
Everard jeune, fabr. d'objets d'albâtre, pass. des Petites-Ecuries, 20.
Everickx, layetier, Temple, 72.
Evette (Félix) NC, ancien juge au tribunal de commerce de la Seine, Petites-Ecuries, 13.
Evette (F.) père et fils NC, commissionn. en charbon de terre, Faub.-St-Martin, 166.
Evilliot, bois de sciage, Chapon, 14.
Evotte, tourneur et menuisier-modeleur, à Batignolles, avenue de Clichy, 150. *
Evrard (Louis), fab. d'albâtre, Bondy, 80.
Evrard (G.), fab. albâtre. Enfants-Rouges, 7.
Evrard, bijoux dorés, Michel-le-Comte, 33.
Evrard (Vor), sculpt. et fabr. de bronzes, St-Louis-Marais, 86. *
Evrard, fab. de chandelles, Popincourt, 97.
Evrard, chapelier, Montmorency, 4.
Evrard et Lebargy, fab. de chapeaux, Blancs-Manteaux, 35.
Evrard (Mme), remise à neuf des dentelles, Hauteville, 30.
Evrard, garniss. en cuivre, Corbeau, 14.
Evrard, mécanicien, Ferdinand-Berthoud, 4.
Evrard, menuisier, Saintonge, 3; ateliers, Popincourt, 61.
Evrard père, menuisier en fauteuils, Charenton, 62. *
Evrard fils, menuisier en faut. Charenton, 63.
Evrard (Ed.), opticien, Blancs-Manteaux, 35. *
Evrard, peintre artiste, Navarin, 10.
Evrard, peintre-vitrier, Hauteville, 30.
Evrard, propriétaire, Guy-la-Brosse, 9.
Evrard, rentier, Victoire, 6.
Evrard, sculpteur-statuaire, Enfer, 91.
Evrard, tailleur, Ste-Anne, 20.
Evrard (L.-F.), verre à vitres, cour des Miracles, 6.
Evrard-Lamotte, habillements d'enfants, Lavandières-Ste-Opportune, 13.
Evrard-Trouillet, gaînier, Montmorency, 33.
Evras, coutelier, pass. Véro-Dodat, 24.
Evrat, bottier, Ste-Anne, 9.
Evrat, chaudronnier, Roi-de-Sicile, 18.
Evrat, professeur au lycée Bonaparte, Pépinière, 11.
Evry (Jules d'), Grenelle-St-Germain, 84.
Exbrayat de Laboriette, contrôleur des contributions directes, Grenelle-St-Germain, 66.
Exelmans (Comtesse Vve), Royale-St-Honoré, 22.
Exelmans (Cte), O. ✻, capitaine de frégate, Isly, 8.
Exmelin, bronzes et curiosités, Nve-St-Augustin, 22.
Exmelin et Arlot aîné, fab. de sellerie, Paradis-Poissonnière, 56.
Expert H., négociant, Université, 2.
Expert (E. Lhuillier, successeur), fab. plumeaux, Petit-Lion, 10.
Extin, fleuriste, St-Denis, 352.
Exupère-Cordonnel, maison spéciale de balances de précision, Rambuteau, 37. *
Eychenne, vins, Corbeau, 27.
Eydoux (Félicité) et Amédée Changey, commissionnaires en marchandises, Gindre, 3.
Eydt, bottier, Pélican, 11.
Eymard (N. et F. ✻) frères, bouchons, Temple, 14.
Eymard (J.-A.), conserves alimentaires, Temple, 14. *
Eymond et Cie, étoffes et apprêts pour fleurs, Bourbon-Villeneuve, 26.
Eynard, C. ✻, général de brigade, Neuve-de-l'Université, 11.
Eynard, hôtel Richelieu, Nve-de-Richelieu, 7.
Eynard, propr., Londres, 27.
Eyraud, avocat, Anjou-Dauphine, 10.
Eynaud, avoc. et recev. de rentes, Beaune, 1.
Eyraud, ébéniste, Greffulhe, 1.
Eyssartier, pharmacien, boul. St-Martin, 41.
Eysen, dessinateur en broderie, Caire, 1.
Ezpeleta (baron F.-X.) ✻, Laffitte, 23.

F.

Faasse aîné, fab. de bijoux dorés, Vieux-Marché-St-Martin, 12.
Faasse, fab. de bijoux dorés, Beaubourg, 105.
Fabart et Tiret, fab. de châles, Fossés-Montmartre, 23.
Fabart, de la maison Fabart et Cie, Faub.-Poissonnière, 72.
Faber (A.-W.), fab. de crayons et art. de bureaux et peinture, boul. de Sébastopol, 12. *
Faber, ébéniste, St-Honoré, 392.
Faber, institution, St-Antoine, 155.
Fabie, grillageur, Marais-St-Germain, 15.
Fabien, rentier, Blanche, 82.
Fabien, rentier, Ferme, 22.
Fabien, vins, Nve-des-Mathurins, 60.
Fabien, vins, Cerisaie, 8.
Fabius, chapeaux et casquettes, Temple, 44.
Fabre (Paul) ✻, avocat au conseil d'État et à la cour de cassation, Guénégaud, 9.
Fabre (J.-A.), ancien avoué, Richelieu, 14.
Fabre, bains Russes, Montmartre, 163.
Fabre, bandagiste, pass. de l'Opéra, galerie du Baromètre, 30.
Fabre, bijout., joaillier, orfèvre, Antin, 13.
Fabre, chaudronnier-mécanicien, Enfer, 64.
Fabre (F.), chiffons et verrerie en gros, Vieille place aux Veaux, 18.
Fabre, économe à la Maternité, Port-Royal, 5.
Fabre, fers en meubles, Montmartre, 27.
Fabre, fontainier, Drouot, 7.
Fabre (E.) fils, grainetier, Helder, 20.
Fabre et Kunemann, ingén.-construct., Faub.-St-Denis, 56, et boul. de Sébastopol, 41.
Fabre jeune, imprimeur-lithographe, Faub.-du-Temple, 129. *
Fabre (G.), loueur de voitures et marchand de chevaux, rue des Champs-Elysées, 7.
Fabre et Hubault, nég.-commissionn., Vendôme, 8.
Fabre, prof. de mathématiques, Ouest, 44.
Fabre, propriétaire, Université, 81.
Fabre, propriétaire, Taitbout, 74.
Fabre, restaurateur, Miroménil, 4.
Fabre (Charles), soies écrues et teintes, St-Denis, 124.
Fabre, tailleur, Faub.-Poissonnière, 13.
Fabre, vins, Notre-Dame-des-Victoires, 28.
Fabre, vins, Larrey, 9.
Fabre, vins en gros, quai de Béthune, 22.
Fabre, vins, Plâtre-du-Temple, 15.
Fabre de Lagrange ✻, chef de bureau au secrétariat général de la préfecture de police.
Fabre de l'Aude (Ctesse), St-Dominique, 28.
Fabre de Parrel ✻, chef de bureau à l'instruction publique, Université, 74.
Fabrège, médecin, Monnaie, 21.
Fabreguette, ancien négociant, Victoire, 94.
Fabreguettes (Eug.) fils et Morra, négoc.-commissionnaires, Faub. St-Denis, 23.
Fabreguettes (F.), professeur de dessin, Paradis-Poissonnière, 47.
Fabri, ancien notaire, Choiseul, 23.
Fabrot, café-estaminet, Feydeau, 24.
Fabrot jeune et Michard, rubans, Nve-St-Eustache, 23.
Fabrot, vins, Fossés-Montmartre, 22.
Fabry, fab. de limes, Ménilmontant, 79. *
Fabulet (Vve), hôt. meub., Quincampoix, 15.
Fabvier, auditeur au conseil d'Etat, Saussaies, 15.
Fabvier (baron), G. O. ✻, général de divi-

Figure 1: Example of the cover page and alphabetical list for the year 1856

The relevance of automatically mining city directories for historical reconstruction has already been argued by several authors (e.g Osborne, N., Hamilton, G. and Macdonald, S. 2014 or Berenbaum, D. et al. (2016). This article reports on the extraction and analysis of the data contained in "Didot-Bottin" covering the period 1839-1922 for Paris, digitized by the Bibliotheque nationale de France. We process more than 27 500 pages to create a database of 4,2 Million entries linking addresses, person mention and activities. The quality of the document analysis process is assessed diachronically and a conservative strategy was chosen in order to populate the database with only information of high confidence. An initial analysis of the data is presented, reporting on the overall statistics of the distribution of professions in Paris and their evolution during more than 80 years, as well a general overview of the diversity of family names through time. Seven case studies corresponding to different streets are briefly compared, showing how information in city directories capture statistically the dynamics of segmentation of the city into functionality differentiated neighborhoods.

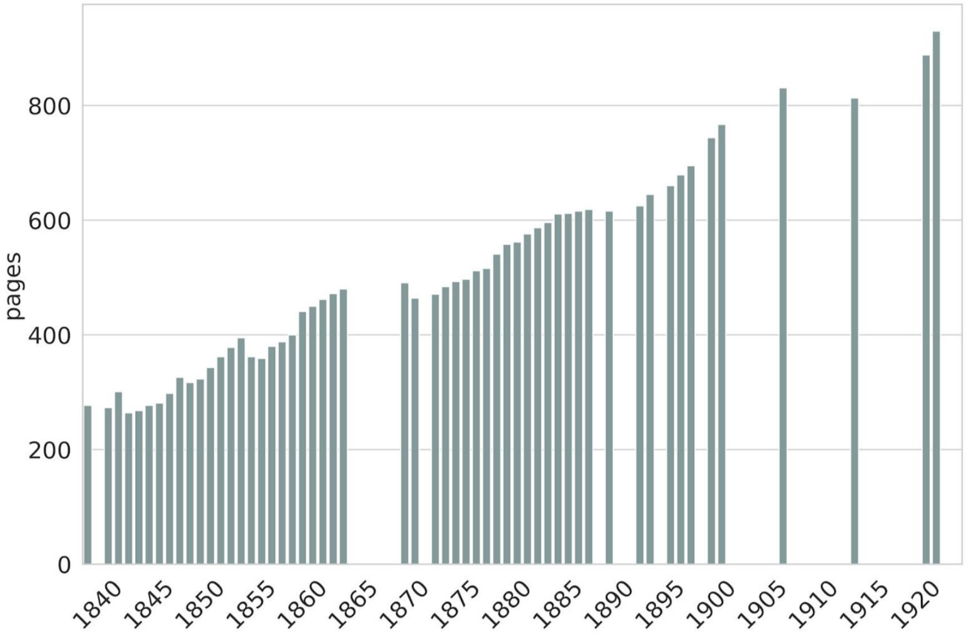
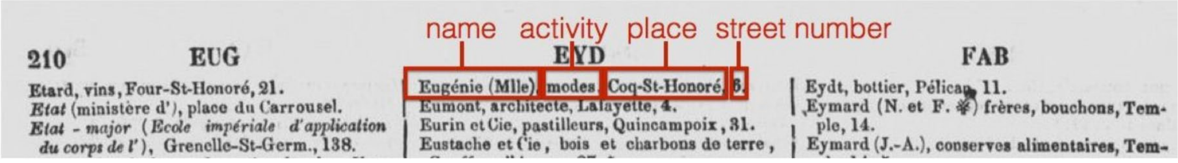


Figure 2: Number of pages per city directory year by year. 29 years are missing in the digital collection used for this study.

2. Method

The document analysed in this article were digitized by Bibliotheque nationale de France and made available online through the Gallica portal. The dataset corresponds to three different published series which are homogeneous in their structure and aims (Bing 1897): Annuaire général du commerce (1839-1856); Annuaire-Didot-Bottin (1857-1907); Annuaire du commerce Didot-Bottin (1914-1922). The documents were associated with an ALTO description containing a structural decomposition of each page into text blocks and lines, associated with a transcription obtained by an Optical Character Recognition (OCR) process. We designed a parsing process converting each line/entry into a record in a database documenting the name, the activity, the place and when relevant the street number (Fig. 3). Only a subset of the entries was successfully parsed (4,2M over 5,6M) and included in the database.



name
activity
place
street number

210 EUG EYD FAB

Etard, vins, Four-St-Honoré, 21.
Etat (ministère d'), place du Carrousel.
Etat - major (*Ecole impériale d'application du corps de l'*), Grenelle-St-Germ., 138.

Eugénie (Mlle), modes, Coq-St-Honoré, 6.
Eumont, architecte, Lalayette, 4.
Eurin et Cie, pastilleurs, Quincampoix, 31.
Eustache et Cie, bois et charbons de terre,

Eydt, bottier, Pélican, 11.
Eymard (N. et F. ✻) frères, bouchons, Temple, 14.
Eymard (J.-A.), conserves alimentaires, Tem-

Figure 3: Structure of an entry in the directories

A general discussion on quality assessment methods of the OCR for the BnF digital collection can be found in (Salah - Moreux - Ragot - Paquet 2015) . In order to assess specifically the OCR quality of the inserted data, 14 pages were randomly picked for the years 1839, 1848, 1856, 1857, 1881, 1907, 1921 and manually controlled.

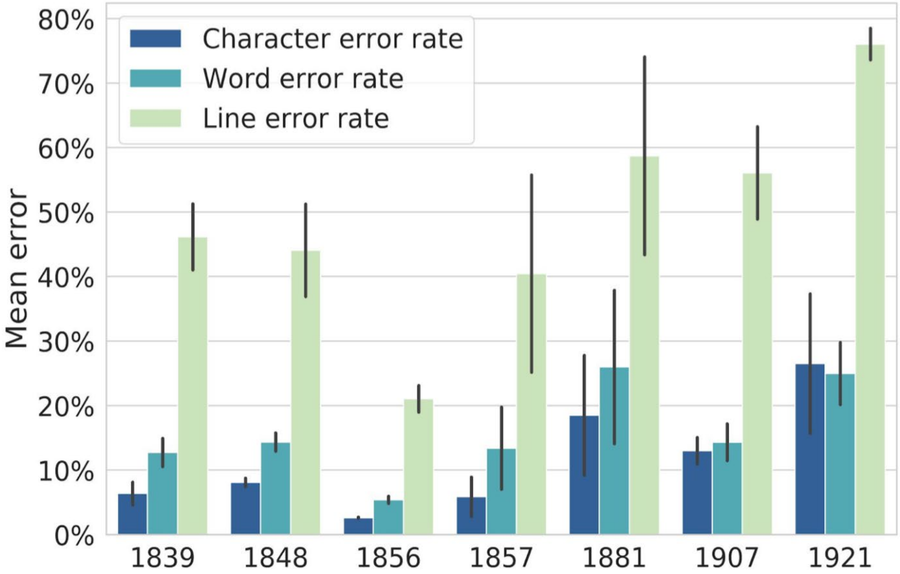
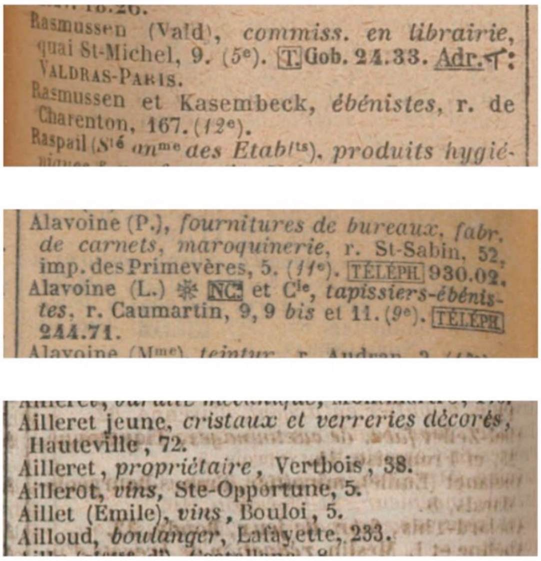


Figure 4: Character error rate, Word error rate and Line error rate for 7 years of the corpus (14 random pages analysed).

On well parsed entries, the mean of the character error is of 2.6% with a standard deviation of 0.1%. The error per line is of 21% with a standard deviation of 2.9% (Fig. 4). The overall transcription quality tends to decrease as sources are more recent. This is essentially due to three factors: (1) the insertion at the beginning of the 20th century of special symbols used for saving textual space but unparsed by most OCR system, (2) the increasing thickness of the volumes leading to the curvature issue during the scanning process making line detection and word identification more difficult (3) the use of continuously thinner paper sheets leading to problems of transparency between the verso and the recto of a page (Fig. 5).



Rasmussen (Vald), *commiss. en librairie*, quai St-Michel, 9. (5e). T. Gob. 24.33. Adr. ☏: VALDRAS-PARIS.
Rasmussen et Kasembeck, *ébénistes*, r. de Charenton, 167. (12e).
Raspail (Sté anme des Etablts), *produits hygié-*

Rasnauss^n (Vald*, commiss. en librairie,
l'lai gt-llichel, 9. (51). MGob. 24.33. :
^LÛIUS-Pahis. ^ Adr,<r
R:.:mussen et Kasembeck, ébénistes, r. de
Raspail -charenton, 167. (12.).

Alavoine (P.), *fournitures de bureaux, fabr. de carnets, maroquinerie*, r. St-Sabin, 52, imp. des Primevères, 5. (11e). TÉLÉPH 930.02.
Alavoine (L.) ❄ NC et Cie, *tapissiers-ébénistes*, r. Caumartin, 9, 9 *bis* et 11. (9e). TÉLÉPH 244.71.

Alavoine (P.), fournitures de bureaux, fakr
de carnets, maroquinerie, r. St-Sabm. __52.
imp.desPrimevères, 5. Uf"). Klépr]930.0s:
Alavoine (L.) !!ml et Cie, tapissiers-é bénis'
tes, r. Caumartin,9,9 bis et 11.(Se).fglgigj

Ailleret jeune, *cristaux et verreries décorés*, Hauteville, 72.
Ailleret, *propriétaire*, Vertbois, 38.
Aillerot, *vins*, Ste-Opportune, 5.
Aillet (Emile), *vins*, Bouloi, 5.
Ailloud, *boulanger*, Lafayette, 233.

Ailleret jeune, cristaux et verreries décorés
llautevîll&, *72-.- . '
Ailleret, propriétaire, VcrUBois", 38.
4iîfert>t, Ste-Gppàrthrie, 5i
Aiilet (E 1 vlle, CotUoi. 5.
Ailloud, •boutiStde^, Eaftf.éwej;29a* "

Figure 5: Examples of problems of scanning (1) symbols, (2) page straightening (3) transparency and corresponding OCR

The entries of the database for seven cases studies were realigned on the Vasserot cadastre digitized and analysed during the ALPAGE project (Noizet-Bove-Costa 2013) and available online. The Vasserot cadastre is giving a full director of addresses in Paris for the period 1810-1837. Paris Open Data covers the structural change due to the Haussmann period and the evolution of the 20th century. Using these two sources, 89,2% of the addresses were successfully realigned for the seven case studies considered.

3. Results

In total 4.2M person mentions were extracted for the period 1839 to 1922. This database could be the starting point of numerous studies, we are only giving here a broad illustration of the content of this dataset and discussing their potential relevance for future research. For instance, the diversity of family names, an indirect proxy of the social effects of urbanization, clearly increases during the 19th century and then stabilizes at the beginning of the 20th century (Fig. 6).

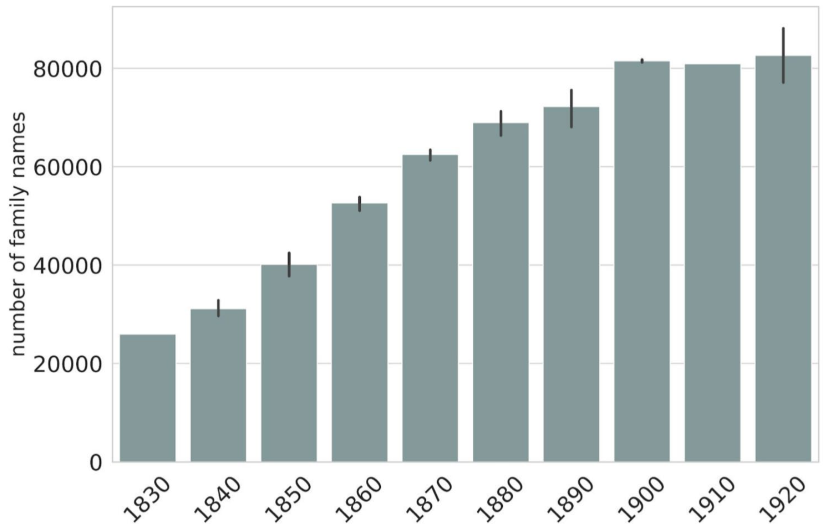


Figure 6: Number of unique family names by year during the entire period. Name diversity keeps increasing during the 19th century and then stabilizes.

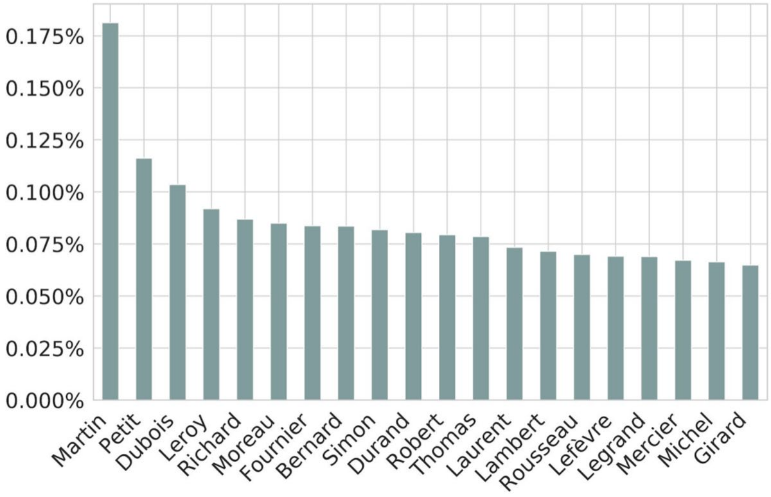


Figure 7: Family names frequency on the period 1839-1922

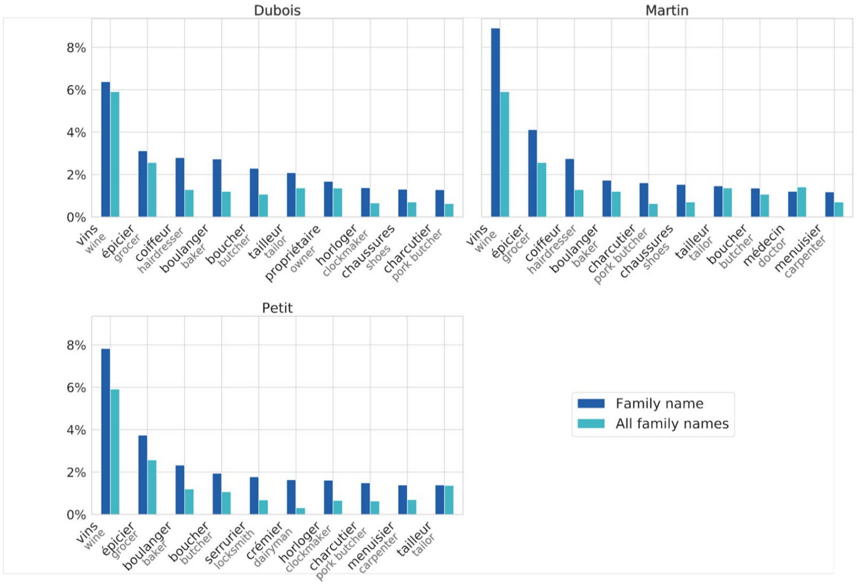


Figure 8: Frequency of trade type by family names “Dubois” ; “Martin” ; “Petit” on the period 1839-1922

If one compares the three most frequent family names “Martin”, “Petit”, “Dubois” (Fig. 7) with the typologies of trading businesses, the sale of wine appears to be, by far, the most commonly practiced activity (Fig. 8). If you encounter a Martin, a Petit or a Dubois at the end of the 19th century, there is a certain chance, he’ll be a wine seller. This result corresponds to a general trend, at the urban scale. In the figures 9 and 10 the dominance of the wine business is 2.5 times more important than the one of grocery stores, the second activity on the list. As confirmed by other studies (Pinol-Garden, 2009), the relative proportion of the wine activity keeps increasing during the 19th century. This is in line with other figures like the wine consumption (one million hectolitre in 1800, 5 million in 1920, about 150-200 litres per person/per year) and the construction of the Bercy wine hall in 1869 (42 hectares on the Seine bank) and their extension in 1910 (Gallet 1939; Thillay-Reynald 2004). Grocery professions, tailor, hairdressers, bakers are equally represented, alongside liberal professions like doctors. Architects, cabinet makers and carpenter are also well represented, a sign of an important building activity on a city scale.

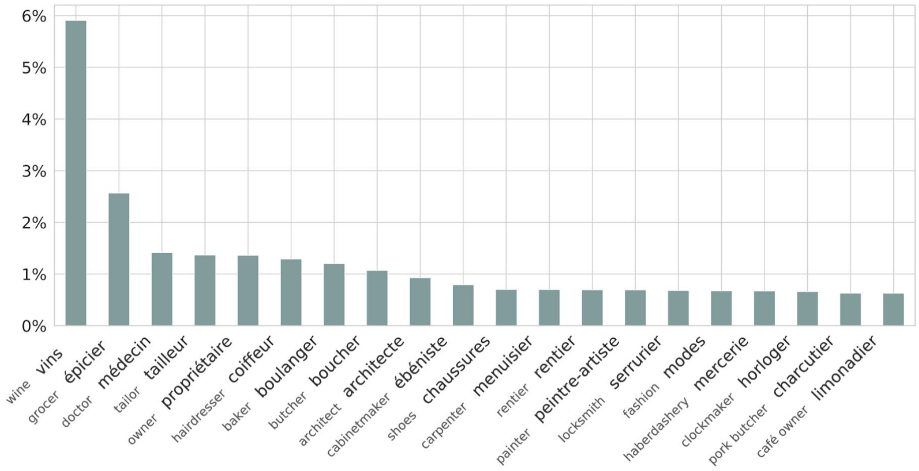


Figure 9: most common trades over the whole period

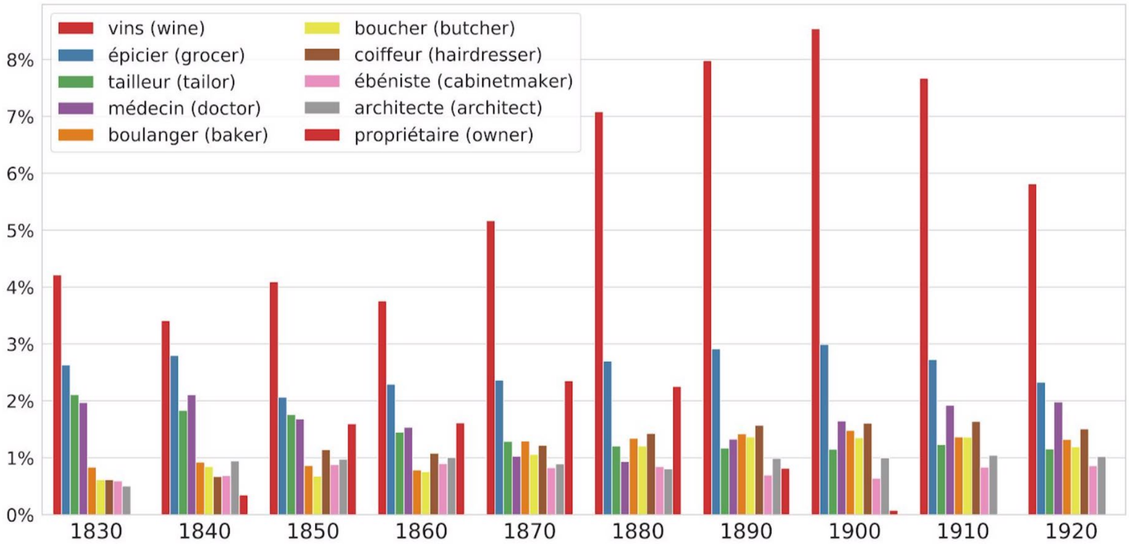


Figure 10: most common trades by decade

The seven chosen case studies correspond to a specific analysis crossing historical information and urban space, focusing on very ancient streets (Rue Montmartre-Rue Saint Denis-Rue Saint Antoine-Rue Saint Jacques) organising a variety of activities over a deep urban “palimpsest” (Corboz 2001) and others, much more recent, with a strong characterization by trade businesses (Rue Richelieu-Rue, Saint Anne) or residential vocation (Boulevard Saint Michel) (Fig. 11). Fig 12 compares the situation of seven cases studies showing the diversification of activities by neighborhood. The ancient roman streets (Rue Saint-Denis, Rue Saint-Antoine, Rue Saint-Jacques) are characterized by the equally distributed presence of services for the population, reflecting a residential activity and a mix of activities acquired over time. Only the mention of the ‘bookbinders’ on the Rue Saint Jacques, highlights the presence of the university district strongly linked to book production. Rue Montmartre, mentioned since the 13th century, includes many more activities than other cases, without clear differentiators, except for the presence of the tailors. Indeed, this axis is the interface to the 'Richelieu District' considered as a place of production of fabrics and clothes. Rue Richelieu and Rue Saint Anne have a concentration of activities linked with fashion (“tailleurs”, “couturière”, “bottiers”). This denotes the original vocation not residential but specific to these trades. The distribution of activities on the Boulevard Saint-Michel constructed after the transformation from Baron Haussmann is significantly different from all the others with higher level of owners and liberal professions (doctors, architects).

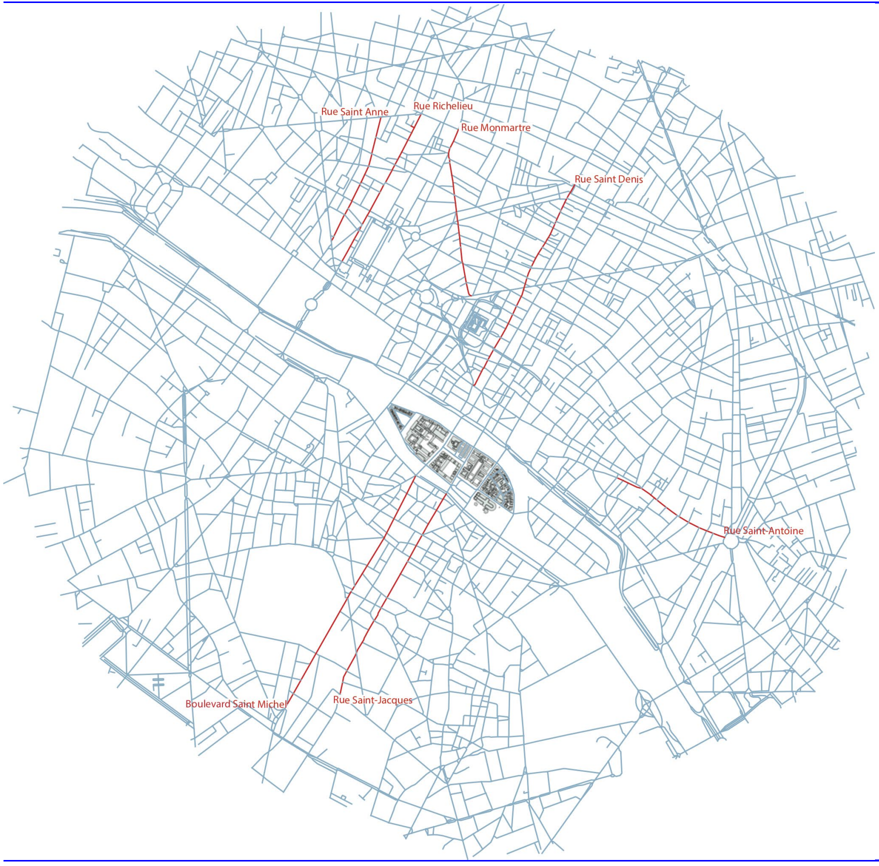


Fig. 11: The seven case studies: *Rue Montmartre; Rue Saint-Denis; Rue Saint-Antoine; Rue Saint-Jacques; Rue Richelieu; Rue Saint Anne; Boulevard Saint-Michel*

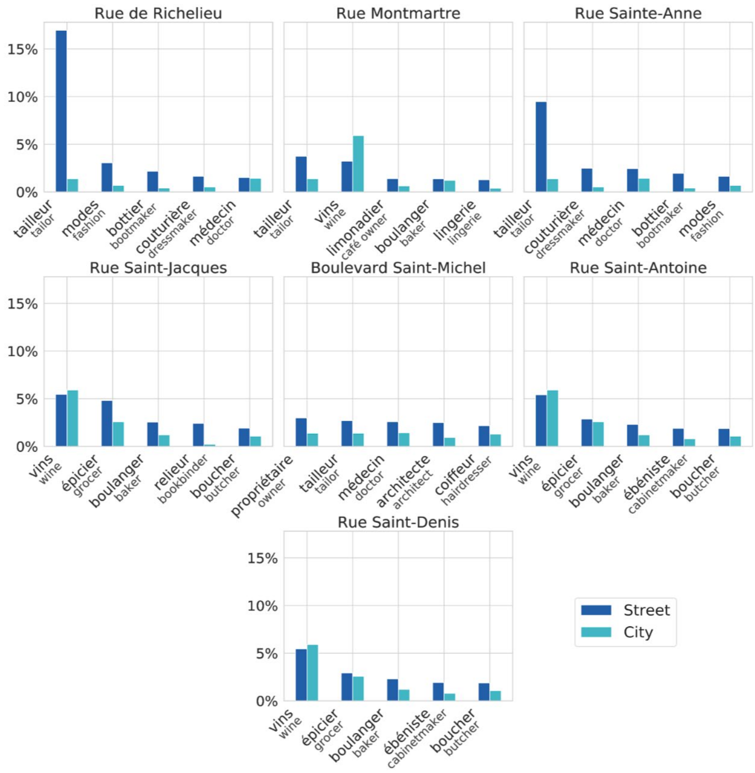


Fig. 12: Street trades (1839-1922)

These preliminary results on Paris hope to demonstrate the potential of city directories to conduct large-scale urban analysis at different level of granularity. The automatic extraction process designed for this article permits to envision to easily conduct similar studies on the population of many other important cities in the world. Provided that the quality of the extraction process is monitored, such kind of massive datasets will open new avenues to study the transformations of the urban structure at different geographical and temporal scales during the ongoing industrialization and other significant societal transformations of the 19th century, connecting these large datasets from the past with the ones of the present.

Appendix A

Bibliography

1. Ben Salah, A., Moreux, J.-P., Ragot, N. and Paquet, T. (2015). OCR performance prediction using cross-OCR alignment. 13th International Conference on Document Analysis and Recognition (ICDAR), pp. 556Ð60
2. Berenbaum, D., Deighan, D., Marlow, T., Lee, A., Frickel, S. and Howison, M. (2016). Mining Spatio-temporal Data on Industrialization from Historical Registries. Journal of Environmental Informatics. d oi:10.3808/jei.201700381 https://arxiv.org/abs/1612.00992
3. Bing, A.-D. A. (1897). Les annuaires parisiens, de Montaigne à Didot, 1500 à 1900. Le Havre.
4. Corboz, A. (2001), Le territoire comme palimpseste et autre essais, Payot, 2001 Gallet, P. (1939). LÕapprovisionnement en vin de Paris. Annales de géographie , vol. 48, no. 274. Paris, pp. 359-368.
5. Noizet, H., Bove, B. and Costa, L. (2013). Paris de parcelles en pixels . A nalyse géomatique de l'espace parisien médiéval et moderne . Paris : Presses universitaires de Vincennes-Comité d'histoire de la Ville de Paris.
6. Osborne, N., Hamilton, G. and Macdonald, S. (2014). Historical Post Office Directory Parser (POD Parser) Software From the Addressing History Project. Journal of Open Research Software , vol. 2, no.1, p. 23.
7. Pinol, J.-L. and Garden, M. (2009). Atlas des Parisiens. De la révolution à nos jours. Paris.
8. Thillay, A. and Reynald, A. (2004). Le grand marché. L'approvisionnement alimentaire de Paris sous l'Ancien Régime. Histoire, économie et société , vol. 23, no. 2 : La société, la guerre, la paix, 1911-1946. pp. 307-308.