

Pop, Pop, Pop...

Les apports des modèles génératifs multimodaux à la démographie historique



www.litislab.fr

Thierry Paquet, Pierrick Tranouez, Clément Chatelain,

Thomas Constum, Nicolas Kempf

Florian Bébin, Théo Larcher, Lucas Préel

en étroite collaboration avec Sandra Brée

HTR en démographie Historique

- C'est aujourd'hui possible
- Obtenir des données de qualité reste difficile
- Des projets longs et exigeants en RH – inter disciplinarité
- Beaucoup d'ingénierie de haut niveau
- Devient difficile de trouver des défis scientifiques TI



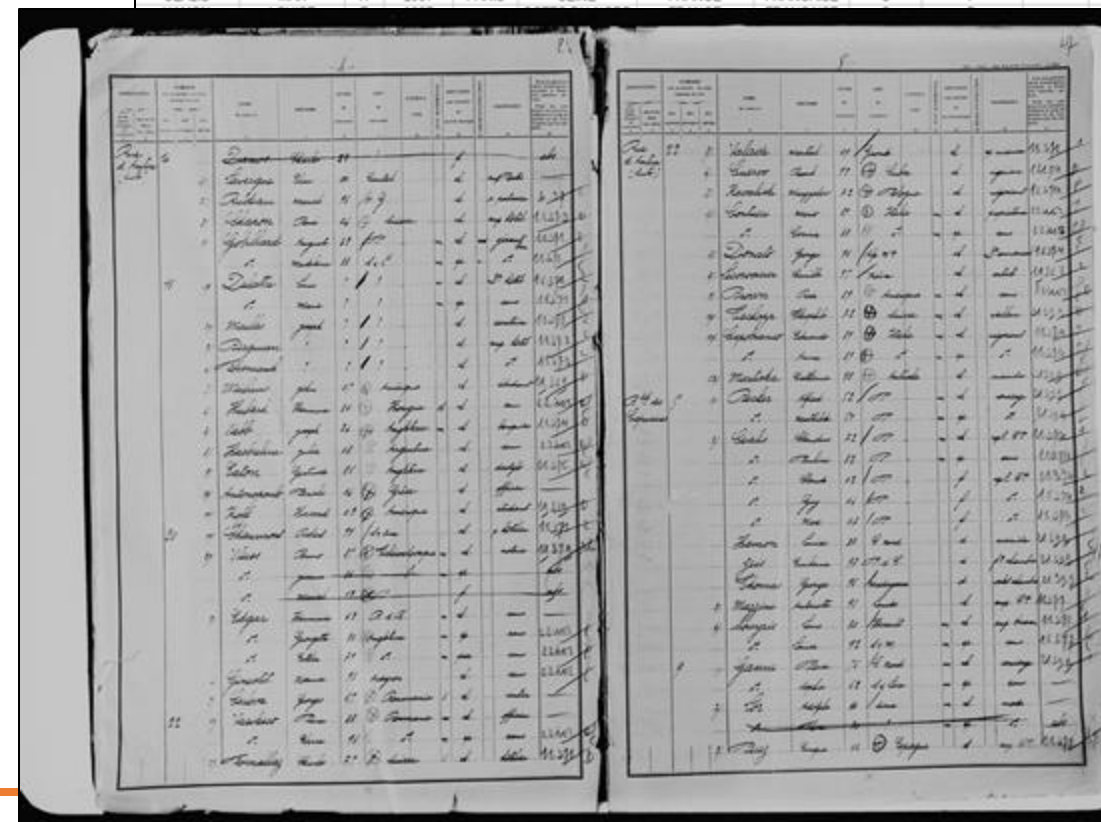
PopStras

- Recensements: POP-CollexPersée (S. Brée)
- Actes de mariages: ExoPOPP-ANR + Région Normandie (S. Brée)
- Registre de population: PopStras-ANR (B. Gérard)

Projet d'Ocrisation de la Population Parisienne

- Mettre en base de données les premiers recensements de Paris (1926, 1931 et 1936).
- 150 000 doubles pages, 30 lignes
- 9 M d'individus

Noms	Prenom	Sexe	Annee de naissance	Ville de naissance	Departement de naissance	Pays de naissance	Nationalite	Etat matrimonial	Situation par rapport au chef de menage	Statut profession	Profession
VALADE	MARTIAL	H	1909		GIRODE	FRANCE	FRANCAISE	C	CH	APP	APPRENTI CUISINIER
CUERVO	RAOUL	H	1893			CUBA	CUBA	C	CH		INGENIEUR
KOVVALISKI	SMEEZYSLOVS	A	1872			POLOGNE	POLOGNE	C	CH		NEGOCIANT
CONTUCCI	MARIO	H	1889			ITALIE	ITALIE	M	CH		PROPRIETAIRE
CONTUCCI	SCORRINAS	F	1887			ITALIE	ITALIE	M	EP		SANS
DONATO	GEORGES	H	1896		ALPES-MARITIMES	FRANCE	FRANCAISE	C	CH	D	DOMESTIQUE COMMERCIAL
LECORVOISIER	SCAMELLES	A	1895		NIEVRE	FRANCE	FRANCAISE	C	CH		ARTISTE
BROWN	ROSE	F	1889			AMERIQUE	AMERIQUE	M	CH		SANS
ESCHOPP	THEOPHILE	H	1872			SUISSE	SUISSE	M	CH		VEILLEUR
CAPOBIANCO	EDUARDO	H	1889			ITALIE	ITALIE	M	CH		NEGOCIANT
CAPOBIANCO	ANNA	F	1889			ITALIE	ITALIE	M	EP		NEGOCIANT
MARTISKA	CATHERINE	F	1898					C	CH		CUISINIERE
BECKER	ALFRED	H	1852	PARIS	SEINE	FRANCE	FRANCAISE	M	CH		CONCIERGE
BECKER	MATHILDE	F	1851					M	EP		CONCIERGE
CEALIS	CLAUDIUS	H	1872	PARIS	SEINE	FRANCE	FRANCAISE	M	CH		REPT CCE
CEALIS	PAULINE	F	1882	PARIS	SEINE	FRANCE	FRANCAISE	M	EP		SANS
CEALIS	CLAUDE	H	1902	PARIS	SEINE	FRANCE	FRANCAISE	C	F		REPT CCE
CEALIS	GUY	H	1904	PARIS	SEINE	FRANCE	FRANCAISE	C	F		REPT CCE
CEALIS	MAX	H	1907	PARIS	SEINE	FRANCE	FRANCAISE	C	F		REPT CCE



FE CHAMBRE
VALET CHAMBRE
EMPLOYE CCE
EMPLOYE BUREAU
SANS
CONCIERGE
SANS
SMEDESS
SMEDESS
EMPLOYE CCE
NEGOCIANT
SANS
SANS
FE CHAMBRE
CUISINIERE
COMPTABLE
COMPTABLE
SANS
PPRENTI EBENISTE
EMPLOYE CCE
MD HOTEL
PATRON FLEURISTE
PATRON FLEURISTE

AD075DP_D2M8_260_0070_warped_right_page_09528.tif

Highest Levenshtein distance n° 0 : 27

Label : Amiot / Pierre / 68 / Imphy Nièvre / M / ch / Jency /

Prediction : Amiot / Lierre / 68 / Impry (Nièr / M. / ch. / cansigons sa dre fenée /



AD075DP_D2M8_260_0070_warped_left_page_09429.tif

Highest Levenshtein distance n° 1 : 21

Label : Chevtchenko / Pierre / 97 / Russe / M / Ch / Serrurier / Romus /

Prediction : Thevtchento / Pierre / 97 / Léégay (luxs Russe / M / ch. / serrurin / Romes /



AD075DP_D2M8_260_0070_warped_left_page_09427.tif

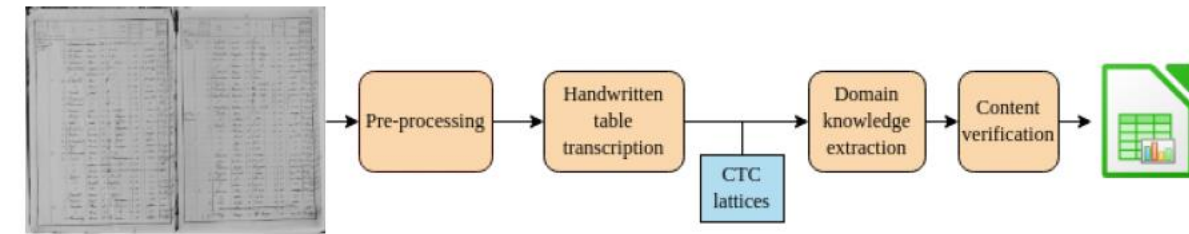
Highest Levenshtein distance n° 2 : 19

Label : Berthouzet / Placide / 96 / Monthey (Suisse) / Suisse / M / ch / E hotel / 18373

Prediction : Berthouzet / Claciae / 96 / Zauttey (Murne / Suisse / M / ch. / Eé hôte / 18373



Projet d'Ocrisation de la Population Parisienne



- Teacher – student training
- Canonisation : métiers – lieux – Abréviations...
- 67 heures de traitement par recensement
- Près de 4 ans de post-correction / validation des données par une équipe de démographes

Model	Dataset	CER (%) validation	WER (%) validation	CER (%) test	WER (%) test
Initial (Model 0)	Generic	6.86	18.66	7.08	19.05
Student 1	Generic	6.07	17.12	6.12	17.12
Student 2	Generic	5.94	16.80	5.97	16.83
Student 2bis	Generic	5.89	16.68	6.02	16.89
Student 3	Generic	5.64	15.98	5.43	15.50
Student 4	Generic	5.01	14.53	4.52	13.57
Student 4 specialized	Mono-writer	2.14	5.24	2.66	6.37

T Constum, N. Kempf, T. Paquet, P. Tranouez, C. Chatelain, S. Brée, and F. Merveille, Recognition and information extraction in historical handwritten tables: toward understanding early 20th century Paris census, 15th IAPR workshop on Document Analysis Systems, pp. 143- 157, La Rochelle, 2022.



Projet d'Ocrisation de la Population Parisienne



[Communities](#) [My dashboard](#) [Log in](#) [Sign up](#)

 IAPR-TC11: Association for Pattern Recognition Technical Committee Number 11

Published March 22, 2022 | Version v1.0

[Dataset](#) [Open](#)

POPP Datasets : Datasets for handwriting recognition from French population census

CONSTUM, Thomas¹  ; KEMPF, Nicolas¹  ; PAQUET, Thierry¹  ; TRANOUEZ, Pierrick¹  ; CHATELAIN, Clément²  ; BREE, Sandra³  ; MERVEILLE, François⁴ 

Show affiliations

POPP datasets

This repository contains 3 datasets created within the POPP project ([Project for the Oceration of the Paris Population Census](#)) for the task of handwriting text recognition. These datasets have been published in *Recognition and information extraction in historical handwritten tables: toward understanding early 20th century Paris census at DAS 2022*.

The 3 datasets are called "Generic dataset", "Belleville", and "Chaussée d'Antin" and contains lines made from the extracted rows of census tables from 1926. Each table in the Paris census contains 30 rows, thus each page in these datasets corresponds to 30 lines.

The structure of each dataset is the following:

- double-pages : images of the double pages

2K

 VIEWS

445

 DOWNLOADS

[Show more details](#)

Versions

Version v1.0	Mar 22, 2022
10.5281/zenodo.6581158	

Cite all versions? You can cite all versions by using the DOI [10.5281/zenodo.6581157](https://doi.org/10.5281/zenodo.6581157). This DOI represents all versions, and will always resolve to the latest one. [Read more](#).

External resources

10/12/2025



Projet d'Ocrisation de la Population Parisienne

- Exposition au musée Carnavalet
- Consultation de la base de données en ligne
- Diffusion de la base de données

Pré-Publication, Document De Travail (Working Paper) ⓘ Année : 2025

POPP. An OCR-Generated Database of the Population Censuses of Paris (1926-36)

Sandra Brée (1) , Victor Gay (2, 3) , Marion Leturcq (4) , Mariia Buidze (5) , Baptiste Coulmont (6) , Yoann Doignon (7) , Thierry Paquet (8) , Thomas Constum (9) , Pierrick Tranouez (8) , François Merveille (10)

10/12/2025



Les gens de Paris, 1926-1936

Interroger par nom ou par adresse dans les recensements de population (1926, 1931 et 1936)



Les listes nominatives de recensement de population des années 1926, 1931 et 1936, numérisées par les Archives de Paris, ont fait l'objet d'un projet innovant mené par le Laboratoire de Recherche Historique Rhône Alpes (CNRS/LARHRA) en collaboration avec le Laboratoire d'Informatique, du Traitement de l'Information et des Systèmes (LITIS) rattaché à l'Université de Rouen Normandie. Il s'agit du projet POPP (Projet d'ocrisation des recensements de la population parisienne). Il a permis de transformer les images en données interrogeables, grâce à un traitement de reconnaissance d'écritures manuscrites utilisant des techniques d'intelligence artificielle (Hand Text Recognition – HTR).

Le LARHRA a partagé une partie des données transcrites dans le cadre du projet POPP. Il est désormais possible d'interroger cette base de plus de 8 millions de noms par nom, prénom, adresse précise (au numéro de rue), et d'accéder ainsi directement à l'image de la page où se trouve la personne recherchée. On peut aussi continuer à interroger par année de recensement, arrondissement, quartier, rue. Les données permettant une exploitation historique (lieu de naissance, profession, composition du ménage, état matrimonial, etc.) ne sont pas disponibles ici.

À noter : les transcriptions effectuées par le traitement automatique ne sont pas toujours fiables, des erreurs subsistent, notamment sur les noms de famille. Le taux de reconnaissance des noms de famille est ici supérieur à 85 %. N'hésitez pas à tester avec différentes orthographes.

[Faire une recherche](#)

DE L'HISTOIRE
DE PARIS CARNAVALET



8 octobre 2025 – 8 février 2026



Les gens
de Paris

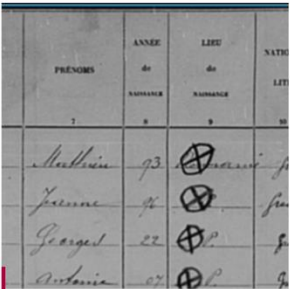
De POPP à ObHisPOP



SCIENTES HUMAINES
& SOCIALES

ActualitésAppels à candidatureLibrairieRevueLettres de l'InSHSAgendaAnnuaireIntranetCNRS.FRFR

À proposRechercheInnovationInternationalTalents



PRÉNOMS	ANNÉE	LIEU	NATURE	LIT.
Mathieu	1873	Paris	Naissance	1
Lucienne	1874	Paris	Naissance	2
Georgette	1875	Paris	Naissance	3
Antonia	1876	Paris	Naissance	4

Accueil > Actualités

ObHisPOP - Observatoire de l'histoire de la population française : grandes bases de données et intelligence artificielle

15 octobre 2024

LA LETTRE HISTOIRE

#ZOOM SUR...

Historienne démographe, chargée de recherche CNRS au [Laboratoire de recherche historique Rhône-Alpes](#) (LARHRA, UMR 5190, CNRS / ENS de Lyon / Université Grenoble Alpes / Université Lumière Lyon 2 / Université Jean Moulin Lyon 3), Sandra Brée s'intéresse à l'histoire démographique des populations, principalement urbaines et banlieusardes, pendant la période 1880-1940. Elle porte le SOSI Observatoire de l'histoire de la population française : grandes bases de données et intelligence artificielle.

A - / A+

Partager ce contenu



Imprimer

22 novembre 2023

Journée annuelle de l'ObHisPop

Observatoire de l'histoire de la population française : grandes bases de données et intelligence artificielle

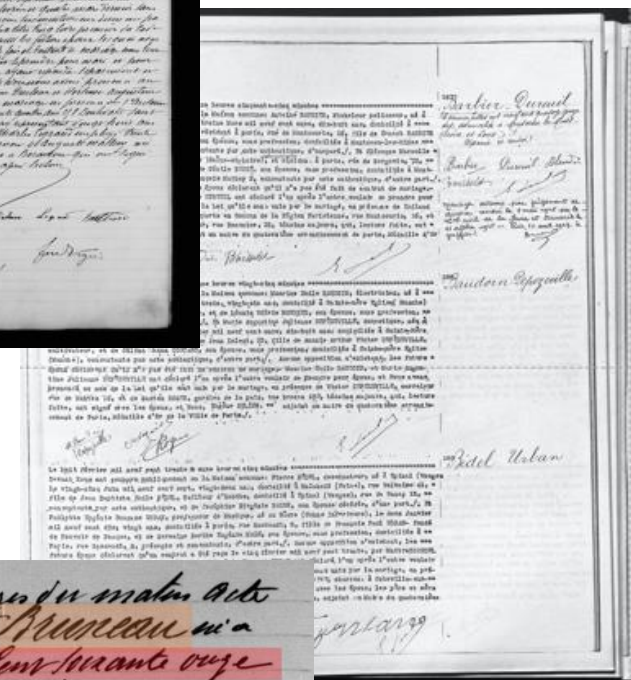
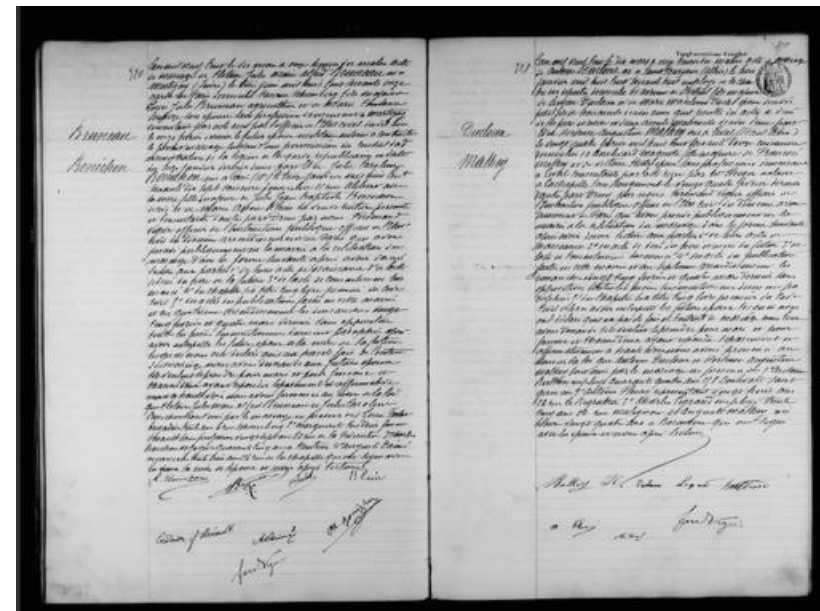
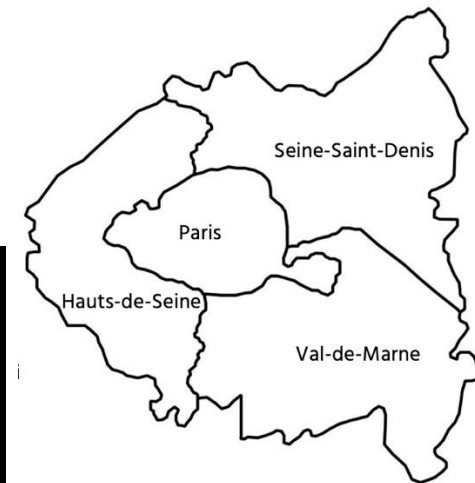
Salle 0.007 (Rez-de-chaussée du bâtiment de recherche nord)
MSH Nord-Campus Condorcet
20, avenue George Sand 93210 La Plaine Saint-Denis
(métro Front Populaire)



Exo-POPP ANR – 2021-2026

Optical Extraction of Handwritten Named Entities for Marriage Certificates for the Population of Paris (1880–1940)

- ✓ 1 M de doubles pages
- ✓ 20 M d'actes
- ✓ 15 à 20 M d'individus



Année 1900 Pour le dix mars, à onze heures du matin acte
de mariage de Clotilde Julie Marie Alfred Bruneau né à
Martigny (Indre) le trois juin mil huit cent soixante onze
garçon de Paris domicilié Casimir Schomburg fils de Jean
Louis Jules Bruneau, agriculteur et de Marie Clotilde

date : 10 Mars 1900
husband's first name : Clotilde
husband's family name : Bruneau
husband's place of birth : Martigny
date of birth : 3 juin mille huit cent soixante onze

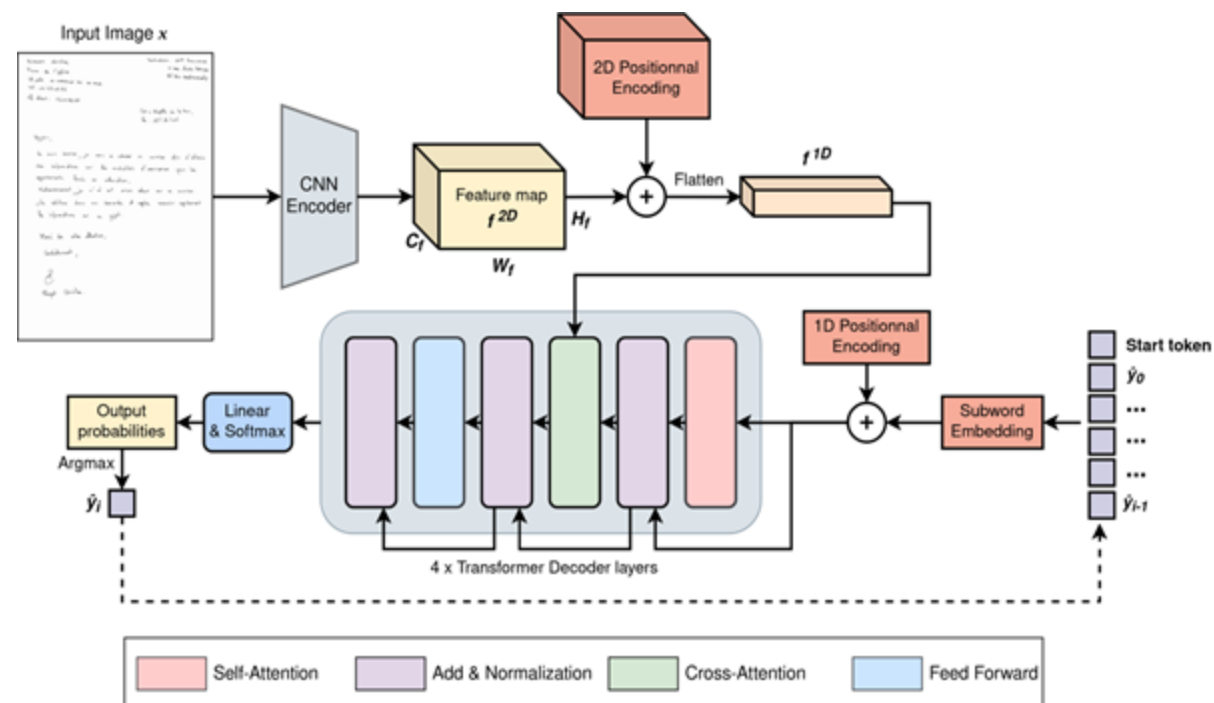
DANIEL

Document Attention Network for Information Extraction and Labelling

DANIEL architecture

DAN inspired conv. encoder, embedding : 1024, 20,6 Mo

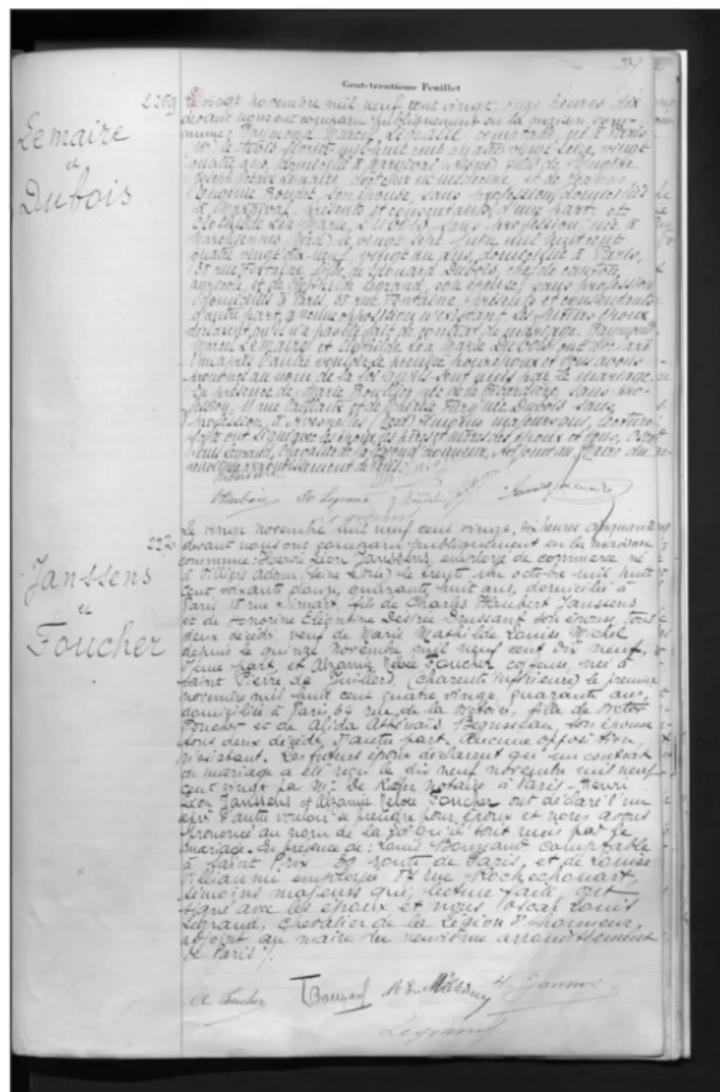
LLM decoder : mBART¹, 134 Mo



D. Coquenat, C. Chatelain and T. Paquet, « DAN: A Segmentation-Free Document Attention Network for Handwritten Document Recognition, » in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 45, no. 7, pp. 8227-8243, 1 July 2023.

T Constum, P Tranouez, T Paquet, DANIEL: A fast Document Attention Network for Information Extraction and Labelling of handwritten documents, IJDAR, January 2025 <https://arxiv.org/pdf/2407.09103>

T Constum, L Preel, T Larcher, T Paquet, P Tranouez, End-to-end information extraction in handwritten documents: Understanding Paris marriage records from 1880 to 1940, ICDAR, pp. 195-214, Springer Nature, 2024, <https://arxiv.org/pdf/2404.19329>



00 ¶
 01 <page>
 02 <record>
 03 <husband-bride-names>
 04 n6
 05 Le maire
 06 et
 07 Dubois
 08 </husband-bride-names>
 09 </paragraph>
 10

¶ @ @ @ _ n 6 ¶ _ Le _ maire _ et _ Dubois ¶ @

Exo-POPP ANR – 2021-2026



CommunitiesMy dashboard

Log inSign up

 IAPR-TC11: Association for Pattern Recognition Technical Committee Number 11There is a [newer version](#) of the record available.

Published April 30, 2024 | Version v1

DatasetOpen

M-POPP datasets: Datasets for full page text recognition and information extraction from French handwritten and printed marriage records

CONSTUM, Thomas (Researcher)^{1,2} ; PREEL, Lucas (Project member)^{1,2} ;
LARCHER, Théo (Project member)^{2,1} ; TRANOUEZ, Pierrick (Researcher)^{1,2} ;
PAQUET, Thierry (Researcher)^{1,2} ; BREE, Sandra (Project leader)^{3,4} 

Show affiliations

M-POPP datasets

This repository contains 2 datasets created within the **EXO-POPP project** ([Optical EXtraction of handwritten named entities for marriage records of the POPulation of Paris](#)) for the task of text recognition and information extraction. These datasets have been published in [End-to-end information extraction in handwritten documents: Understanding Paris marriage records from 1880 to 1940](#) [1] at ICIDAR 2024

999
VIEWS

206
DOWNLOADS

Show more details

Versions

Version v3	Jul 13, 2024
10.5281/zenodo.12737132	
Version v2	May 25, 2024
10.5281/zenodo.11296970	
Version v1	Apr 30, 2024
10.5281/zenodo.10980636	

DANIEL

Document Attention Network for Information extraction and Labelling

Recognition performance

M-POPP

Méthode	CER	WER	LOER	mAP _{CER}
DAN - HTR (Constum <i>et al.</i> , 2024b)	7.21	16.42	5.35	83.03
DAN - HTR+IE (Constum <i>et al.</i> , 2024b)	6.52	14.80	3.79	86.29
DANIEL - stratégie A	8.35	17.36	5.12	80.98
DANIEL - stratégie B	6.61	15.33	6.01	84.41
DANIEL - stratégie C	5.72	14.08	1.34	89.28
DANIEL - stratégie D	14.28	25.47	3.56	67.10

RIMES 2009

Méthode	CER	WER	LOER	mAP _{CER}
DAN (Coquenet <i>et al.</i> , 2023a)	4.54	11.85	3.82	93.74
Faster-DAN (Coquenet <i>et al.</i> , 2023c)	6.38	13.69	4.48	91.00
DANIEL - stratégie A	6.01	11.40	5.50	90.13
DANIEL - stratégie B	5.81	11.24	4.60	91.99
DANIEL - stratégie C	5.80	11.22	6.57	90.21
DANIEL - stratégie D	6.43	12.04	5.26	89.44

IAM

Méthode	CER	WER
OrigamiNet (Yousef et Bishop, 2020)	4.7	-
VAN (Coquenet <i>et al.</i> , 2023b)	4.5	14.6
DAN (Tarride <i>et al.</i> , 2023)	4.3	13.66
Dessurt (Davis <i>et al.</i> , 2022)	4.8	10.2
DANIEL - stratégie A	5.02	11.95
DANIEL - stratégie B	4.38	10.89
DANIEL - stratégie C	4.75	11.44
DANIEL - stratégie D	5.98	13.98

DANIEL

Document Attention Network for Information extraction and Labelling

Named Entity Recognition

IAM

Méthode	F1 (RWTH)	F1 (<i>custom</i>)	Type
Toledo <i>et al.</i> (2019)	14.9	18.0	Séquentiel
Rowtula et Krishnan (2018)	32.3	30.3	Séquentiel
Tüselmann <i>et al.</i> (2021)	52.0	53.6	Séquentiel
Dessurt (Davis <i>et al.</i> , 2022)	40.4	48.5	De bout en bout
DAN (Tarride <i>et al.</i> , 2023)	31.3	-	De bout en bout
DANIEL - stratégie A	46.35	54.83	De bout en bout
DANIEL - stratégie B	48.86	58.29	De bout en bout
DANIEL - stratégie C	50.06	57.55	De bout en bout
DANIEL - stratégie D	41.79	53.25	De bout en bout

M-POPP

Méthode	F1
DAN (Constum <i>et al.</i> , 2024b)	76.37
DANIEL - stratégie A	71.71
DANIEL - stratégie B	74.20
DANIEL - stratégie C	76.37
DANIEL - stratégie D	59.64

DANIEL

Document Attention Network for Information extraction and Labelling

Speed

Handwriting recognition

Méthode	RIMES 2009	READ 2016	IAM	M-POPP
DAN (Coquen et al., 2023a)	4.94	3.84	2.85	15.5
Faster-DAN (Coquen et al., 2023c)	0.99	0.84	-	-
Dessurt (Davis et al., 2022)	-	-	1.48	-
DANIEL	0.77	0.80	0.37	2.66

with NER extraction

Méthode	IAM NER	M-POPP NER
DAN (Coquen et al., 2023a)	3.24	17.94
DANIEL	0.51	3.16

PIVAN

Plateforme d'Indexation et de Valorisation d'Archives Numérisées

Florian Bébin

118 NER categories

NER : name, surname, location, date, occupation..

LINK: Husband's father, Husband's occupation...

T. Constum, F. Bébin, P. Tranouez, T. Paquet, *Pivan: A Web-platform for Document Annotation* – Journal of the Archiving Conference, Vol. 20, n°1, pp. 53-56, Society for Imaging Science and Technology, 2023.

The screenshot displays the PIVAN - ANNOTATEUR web interface. At the top, there's a header with the PIVAN logo and navigation links. The main area shows a document titled 'AD092EC_5MI_CHT_03_0155' with a sidebar on the left containing a list of documents. The document text is displayed in the center, with various NER categories highlighted in colored boxes. A sidebar on the right lists the categories used: MARI, EPOUSE, TÉMOIN, PÈRE, MÈRE, NAISSANCE, RÉSIDENCE, DÉCÈS, AGE, PROFESSION, PRENOM, NOM, DÉPARTEMENT, and VILLE. The document text is a marriage record from 1862, mentioning Auguste, Flécheau, and Juliette, Marie, Couronne, Domestic, Demeurant de fait en cette commune rue du Ponceau quarante-neuf, chez son oncle paternel le Sieur Victor, Couronne, et de droit chez son père à Charné (Sarthe) fille mineure et légitime de Auguste, Honoré, Couronne, jardinier au dit Charné et de Joséphine, Léonide, Lecomte, son épouse, décédée, âgée de dix huit ans, née au dit Charné, arrondissement de Mamers (Sarthe) le vingt deux janvier mil huit cent soixante deux. Le père a déclaré donner consentement au mariage de sa fille avec le Sieur Auguste, Flécheau, par acte passé pardevant son maître. Le maître, notaire à Charné en présence de témoins le trente Août mil huit cent quatre vingts. Les comparants nous ont requis de procéder à la célébration de leur mariage dont les publications ont été faites en cette commune les dimanche douze et dix neuf Septembre dernier sans opposition ainsi qu'il appert le registre des publications de mariage qui nous a été représenté, et en la commune de Charné (Sarthe) les mêmes dimanches également sans opposition, ainsi qu'il est constaté par un certificat délivré par le maire de la dite commune en date du vingt quatre du dit mois de Septembre. Les futurs et les personnes présentées pour autoriser le mariage interpellés par nous en réactions de la Loi du dix Juillet mil huit cent cinquante, nous ont déclaré qu'il n'a pas été fait de contrat de mariage, Les pièces déposées pour être conservées au dossier dont les acte de naissance de chacun des futurs, d'acte de décès de la mère de la future et acte de consentement au mariage ainsi que le certificat de publications et de non opposition Jrovisé, toutes

Exo-POPP

ANR – 2021-2026

Produire des données

**Années 0 ...
et finalement toutes les années**

- ✓ 118 NER categories
- ✓ 1 M de doubles pages
- ✓ 20 M d'actes
- ✓ 15 à 20 M d'individus

- Détecter les données incohérentes au moment de la production
 - ✓ Champs manquants
 - ✓ Champs dupliqués
 - ✓ Champs incertains
- Mettre en place des correctifs automatiques quand c'est possible
- Détecter des incohérences en SHS
 - ✓ Mettre les erreurs en évidence et leur fréquence
 - ✓ Interaction SHS-TI analyser les causes
 - ✓ Trouver des correctifs quand c'est possible
- Sur Exo-POPP 3 mois d'interaction
 - ✓ La plupart des erreurs détectées trouvent une explication logique en retournant à la source
 - ✓ Trouvent un correctif
 - ✓ DANIEL a fait des erreurs cohérentes !

130 pages annotées suffisent pour le spécialiser sur les matrices cadastrales!!!

[illegible]

DANIEL sur le terrain...

Conversion des tableurs dans un format texte compréhensible par la machine

Symboles “\$” et “^” repris de POPP pour indiquer des éléments au dessus ou en dessous de la ligne

Symboles “_” avant et après un mot pour indiquer que ce mot est barré

Utilisation de “id” pour indiquer les dittos

- Klein - - Jean - - Pierre - - Arsdorf -
- 155 -

A; - 355 -;- 2448 -;- 793 - ;in Fresselbourn;;;1924;154

Art. Klein Nicolas - Bilsdorf - 1908
156 - Arsdorf - 1911
Bilsdorf 1934

A;650;1401;427;auf der Heid;;;;
;653;3152;933;;;;

; - 253 - ;;;Arsdorf;1907;62;1908;156

; - 254 - ;;;;id;id;1911;407

Total de

; - 253 -;- 3341 -;- 957 - ;Arsdorf;1908;156;id;id

Total de

id

DANIEL sur le terrain...

Apport de données synthétiques

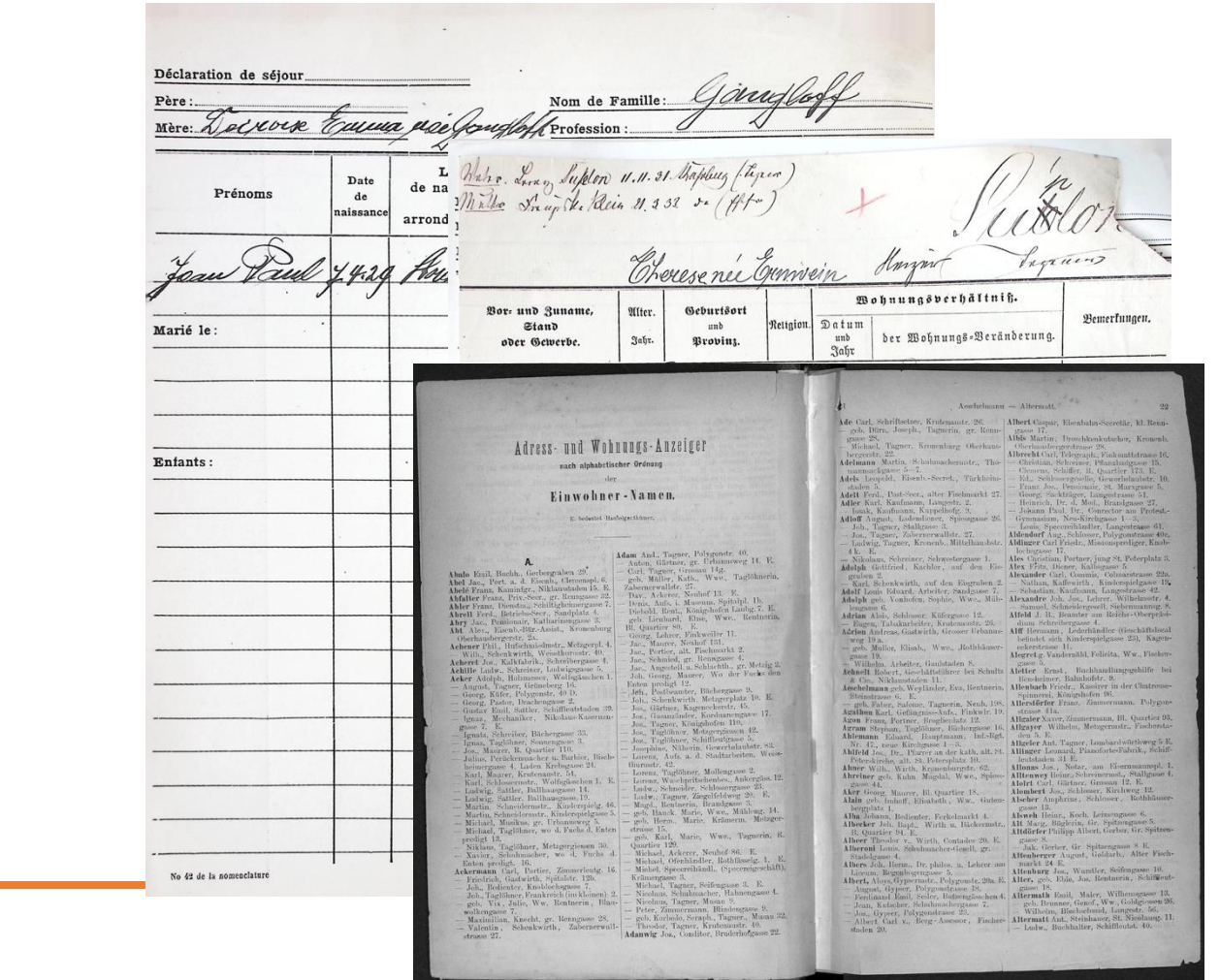
	CER valid	WER urbia	CER test	WER urbia test
1	2,37 %	4,62 %	6,82 %	9,16 %
2	4,32 %	7,77 %	5,77 %	8,58 %
3	4,44 %	6,94 %	5,63 %	7,47 %
4	2,89 %	7,20 %	8,09 %	9,85 %
5	5,12 %	8,72 %	5,08 %	7,69 %
Moyenne	3,83 %	7,05 %	6,28 %	8,55 %
Ecart-type	1,03 %	7,54 %	1,07 %	8,43 %

Article.	NOM.			PRÉNOM.		NOM de la femme ou autre indication distinctive.		PROFESSION.		DEMEURE.		ANNÉE de la mutation.	
	Lambert			Joseph				industriel		Derenbach		1772	

SECTION.	NUMÉRO des PARCELLES.	NUMÉRO supplémentaire.	Cote du Croquis d'arpentage.	NOUVEAU NUMÉRO.	LIEU-DIT.	NATURE de CULTURE.	CONTENANCE.			CLASSES.	REVENU NET.		REVENI DES PARCELLES MUTÉES.				
											TIRÉE		REPORTÉE				
											en l'année de l'article		en l'année de l'article				
1.	2.	3.	4.	5.	6.	7.	8.			9.	10.		11.	12.	13.	14.	
D	67	4570	80		Vor Kuemes	lab.	73	79	81	5	370	76		4	1853	15067	
	12	122	691		Buchholzwies	jardin	36	10	54	4	304	87			1891	18952	
	2109		76		Beim Koppelbesch	haie		75	80	2	987	19			1905	187	
	5092		15		Rue Biergenkraz	brouss.		70	57	9	889	42	3905	1858	789		
	1611	640	91		Im Hinterslen Loch	haie		99	21	3	220	12			1865	5072	
	11	4535	16		Rue Simone Veil	haie		93	16	6	983	75			1798	44905	
	780	490			Auf Dem Bredendriesch	ble		98	63	5	397	56	32	1830	99876		
	674	659	811		Rohlber	brouss.	35	79	34	5	302	55			1714	352450	
	8990	9508			Dellwies	lab.		56	46	7	675	33	36803	1893	844		
	5484	92			Schnokelach	ble	57	16	49	2	150	24	614682	1890	168		
	779		803		An Der Tomm	jardin		77	26	7	121	10	1823	1802	517		
	3737		207			brouss.		37	44	3	473	19		1899	3359		
	40		59		Rue De Wecker	jardin		51	89	8	288	54			1746	78630	
	2362	942	65		Auf Zaemer	ble	35	37	59	3	498	69	1779	1902	75774		
	898	98	600			palure		32	93	8	561	80	1734	9	1770	609	
	768	282	409			brouss.		88	82	6	925	64	78	1759	197		
	9136	51			An De Linnefelder	haie		78	10	5	272	18	83	1885	124		
	8206	970	732		Auf Baumbuscherwies	jardin	41	53	45	1	106	67	9442	1719	5076		
	40		642		In Der Achl	haie	39	72	83	0	882	55			1873	238737	
	2313	6666	98			jardin		26	72	6	701	84			1711	23926	
	591		525		Rue De Syren	lab.	26	94	83	0	785	35	36676	1880	9799		
	69	8885	725		Rue Du Fosse	lab.		98	71	6	543	89			1818	462966	
	215	3261	335		Um Kellenhouschl	ble		93	70	6	415	42			1713	500	
	990		58		Schloss	ble		21	96	5	353	42	6498	1831	834		
	4139	39	303			palure	50	87	63	0	535	89			1886	238981	
	134	552	57		Rue De La Gare	lab.		97	74	1	624	89	1757	1714	933731		
	59	736	318		Rue Jean Arp	palure		21	82	8	409	73			1774	399	
	9184	6181			Allée Du Carmel	palure		29	88	6	649	34			1872	1622	
	64		44		An Der Schamheck	ble		24	77	0	813	85			1760	470938	
	4528	574	62			palure		73	19	4	127	63			1830	409	
	99	41	37		In Lag	lab.		29	41	0	276	31			1784	7170	
	414	7338			In Der Koenigswies	lab.	69	67	64	5	258	67	1788	36917	1796	257	
	973	75	362		In Weierbach	jardin	14	31	93	2	590	61			1874	186152	
	3840	645	28		Beim Neuen Weiher	ble	33	84	52	0	116	94			1871	6707	
	693	839	944		Rue Des Marlyrs	jardin	72	44	17	4	283	73	1748	355	1709	5149	
	1470	5984			Spalzdriesch	haie	43	65	72	4	730	54			1782	4563	
	77	933			Enzbongert	ble	15	24	14	0	584	50			1815	9586	
	672		80		Im Dueren Brill	jardin		98	50	5	710	63			1731	13882	
	77	748	199			lab.		73	28	8	441	57			1810	509574	
	91		97		Im Aeltchen	ble	19	23	31	3	802	92		867126	1787	5667	
471	41			Rue De Slavelot	ble		28	19	5	395	31			1741	906		
3204	9741	88			palure		14	20	7	576	41		4	1885	6415		

MODÈLE N° 1.										PAGE 54				
Article	NOM.		PRÉNOM.	NOM de la femme ou autre indication distinctive.	PROFESSION.	DEMEURE.	ANNÉE de la mutation.							
	Franck							Jean Pierre	Schaaf	cultivateur	Kaltbühl	1924		
60														
SECTION. NOMBRE des PARCELLES	NOMBRE supplémentaire.	Classe de Censité d'origine.	NOMBRE NUTRÉO.	LIEU-DIT.	NATURE de CULTURE.	CONTESANCE			CLASSES.	REVENU NET.	RENTES DES PARCELLES MUTÉES.			
						7.	8.	9.			10.	TIRÉE		REPORTÉE
1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	
B	463	452	259	✓	in Kithier	lab				4	5.60		1928	60
	463	454	id	✓		pré				4	3.11		id	id
	463	452	1441		Waldkilling	man. exp.	25.50	40		47.50			1929	463
	655				zu Ritz	pré	10.50	3		14.14				
	752	3008	1242	✓	in Lamsbach	lab	55.00	3		10.58				
	752	3000	id	✓	in Lamsbach	lab	55.00	4		5.56				
	811	485	263	✓	in Lamsbach	lab	50.00	4		6.11				
	1083	1192	575	✓	in Lamsbach	bois	10.50	3		1.12				
	1087			✓		lab	10.50	3		6.12				
	1219			✓	in Brachbach	lab	10.50	3		3.11				
	1515	2516	1261	✓	Waldkilling	lab	25.50	4		5.14				
	1087	1640	1082	✓	Waldkilling	bois	55.00	3		3.10			1914	200
	1199			✓	auf dem Bach	lab	20.00	4		5.32				
	1575	3205	1261	✓	auf dem Berg	lab	25.50	3		9.04				
	1581	3209	id	✓		lab	20.00	4		5.32				
	1503	3258	1261	✓	am spitzen Binsbaum	lab	10.50	4		6.01				
	1547	575	274	✓	vor Hahn	pré	10.50	4		5.11				
	1550	579	id	✓		lab	10.50	5		4.39				
	1561	1001	579	✓	auf der Kugelbach	pré	11.40	4		5.53			1922	161
	1561	1002	id	✓		pré	10.50	4		3.34			2	2
	1562	574	274	✓		lab	55.00	3-1-2-4		9.10				
	1600	3160	1270	✓	Waldkilling	pré	12.50	4		3.22				
					Total		44.00			11.14				
	721	101	99		in Lamsbach	pré	10.50	3		3.02	1907	20		
					Total de 1908		5.21.43			19.25				
	463	457	259		in Kithier	lab	20.00	4		5.18	1908	90	1928	60
					Total de 1909		5.41.02			19.77				
	381	4332	679		Waldkilling	bois	55	0		38	1910	38		
	1601	641	376		Waldkilling	lab	20.00	4		4.24		64		
					Total de 1911		5.41.48			30.22				
	1083	1192	575		in Lamsbach	bois	18.00	3		2.08	1913	197		
	1084	750	359			lab	20.00	4		4.32				
					Total de 1914		5.49.18			20.87.5				
					1915		5.75.53			20.5.03				
	658	20	53		zu Ritz	pré	10.50	3		4.32	1916	90		
	658	25				pré	10.00	3		4.20		2	2	
					Total de 1917		5.77.55			21.3.43				
	307				Waldkilling	pré	10.50	3		2.17	1918	13		
	308					lab	25.50	3		1.41		2		
	312	2937	1208			lab	10.50	2		11.52		11		
					Total de 1919		6.46.48			2.47.28				
	1602				Waldkilling	lab	25.50	4		5.50	1920	108		
					Total de 1921		6.47.27			2.47.48				

00



- # ANR – 2025-2030
- ## POP-Stras
- Univ. Strasbourg – SAGE – Bénédicte Gérard
Univ. Rouen – LITIS – UMR IDEES
- Le fichier domiciliaire de la ville de Strasbourg (1870-1940)
 - ✓ Situation des ménages et leurs mouvements
 - ✓ Mis en place sous le régime Prussien
 - ✓ Repris par l'administration Française
 - Compréhension de documents semi-structurés multilingues par un système de questions/réponses visuelles.
 - Exploitation de sources connexes : les annuaires

ANR – 2025-2030
POP-Stras

Univ. Strasbourg – SAGE – Bénédicte Gérard
Univ. Rouen – LITIS – UMR IDEES

- **Le traitement des annuaires**

1 Abby (OCR)

2 Post-traitement des sorties

3 Extraction avec un LLM

31

Balthes — Barthel

32

Barthel

33

Bartel

34

Bauer

Balthes Carl, Locomotivführer, Alterweimmarkt 10.
Balthes Carl, Professor, Krutenaustrasse 2.
— Daniel, Rentner, Langestrasse 93.
— Michael, Wirth, Polygonstrasse 61. E.
Baltzinger geb. Rothbach, Anna, Rentnerin, Magdalengasse 6.
Baltzi, Joh., Tagner, Lingolsheimerstr. 19.
— Georg, Brauknecht, Königshofen 44.
— Georg, Taglöhner, Kinderspielg. 10.
Balneo Heor, Hutmacher, Langestrasse 10.
Balzer Georg, Bremser, Kronenb., Mittelhandelsplatz 2b.
Balchner Ludwig, Ackerer, Neuhof 93. E.
Balzl Georg, Dienstmann, Thomasstadten 91.
Bambis Dorethe, Wirthin, Johannesstader 5.
— Emil, Schreiner, gr. Stadelgasse 6.
— geb. Hoh, Ww., Salome, Gärtnerin, R. Quartier 7. E.
Bammes Gottfried, Rentner, Neuer Kirchhofen 31.
Bandolzy Carl, Putschaff, Nikolausstad. 15.
Bangardt geb. Gotzel, Marie, Ww., Königshofen, Eckoldshausstrasse 3. E.
Bangard Michel, Schreiner- und Glasermeister, Kinderspielen 91.
Banzatz Florenz, Tagner, Musau 6. E.
— Laurent, Gastwirth, Ludwigsgasse 24.
— Theobald, Tagner, Klebsauweg 3. E.
— Valentin, Wirth, R. Quartier 7. E.
Banhard Sebastian, Gypsfigurenmacher, Spitalgasse 13.
Bankhauser Aug., Gärtner, Ziegelweg 29.
Banschuh Anton, Handlungs-Commis, Finkmattstrasse 15.
Banshaf Julius, Dienstmann, Stampfgasse 1.
Bantelm Albert, Sattler, Artilleriewallstrasse 1.
Banzet Ludw., Schreiner, Langestr. 19.
Bapst Philipp, Specereihändler, Hellebardengässchen 2.
Barrack August, Prof. Dr., Oberbibliothekar der Schloss-Antiquarisch-Schlossplatz 2.
Baraniki Aug., Postsecret., Allerheiligeng. 43b.
Barasky Wilh., Eisenbahn-Sekretär, Königshofen 43b.
Barbaness Catharina, Rentnerin, Schlachthausplat 3.
— Georg, Gastwirth, Schleifhausplatz 33.
— Louise, Rentnerin, Gewerkslaube 9.
Barbas geb. Heiser, Magd., Ww., Rentner, Friedhofen 23.
Barbier Bernhard, Octroi-Beamter, Lindenfelsergasse 10.
— Georg, Pfisterer, Bruderhofgasse 39.
— geb. Kientz, Magdalena, Ww., Tagnern 19.
— geb. Herrenberger, Maria Anna, Schenken 19.
— Nicolaus, Pflasterer, Blumenstaden 4.
Barchewitz Julie, Ww., Schildgasse 5.
Bardel Osc., Octr.-Einnehm., Schlachthausplat 2.
Bardes Ludwig, Metzger, Ballhausgasse 2.

Bardol Anton, Schreiner, Finkweiler 3.
— Johann, Ackermann, Königshofen 34. E.
Bardotet Adolph, Maschinen-Techniker, Altkrutenaustrasse 40.
Bareis Aug., Würstler, Schiffleutstader 22. E.
— Barmine geb. Barenne, Maria v., Ww., Renbfuss Friedr., Postschaffner, g. Rentholz Wilhelm, Tagner, R. Quartier 2.
Barkhausen Gustav, Kleiderhändler, Ver-

Bartel Theob., Tagner, Neuhoft 14.
Barthme August, Bäckermeister, Zimmergasse 20.
— Fritz, Tagner, Gr. Spitzgasse 4.
— Gust. Theob., Müchl-, Lumpenhändl. 143.
— Xaver, Maurer, Planchholz 12.

Bart geb. Grünigler, Magd., Lumpenhändl. 143.
Barten Sybilgasse 21.
Battermann Carl, Ofenmacher, Metzgeressen 30. E.
Battezzoli Fridolin, Maler, Büchersgasse 12.
Batz, Fritz, Schlossplatz 2.

<column>
Balthes Carl, Locomotivführer, Alterweimmarkt 10.
Baltzer Carl, Professor, Krutenaustrasse 2.
— Daniel, Rentner, Langestrasse 93.
— Michael, Wirth, Polygonstrasse 61. E.
Baltzinger geb. Rothbach, Anna, Rentnerin, Magdalengasse 6.
Baltzli, Joh., Tagner, Lingolsheimerstr.
Baltzli, Joh., Tagner, Lingolsheimerstr.
— Georg, Brauknecht, Königshofen 44.
— Georg, Taglöhner, Kinderspielg. 16.
Balucco Heinr., Hutmacher, Langestrasse 1
Balzer Georg, Bremser, Kronenb., Mittel
Balzinger Ludwig, Ackerer, Neuhoft 93. E.
Balzli Georg, Dienstmann, Thomasstad
Bambis Dorethe, Wirthin, Johannesstader
— Emil, Schreiner, gr. Stadelgasse 6
— geb. Hoh, Ww., Salome, Gärtnerin, gr.
Rennegasse 12. E.
Bammes Gottfried, Rentner, Neuer Kirche
Bangardt geb. Gotzel, Marie, Ww., Könige
Bangard Michel, Schreiner- und Glaserme
Bangratz Florenz, Tagner, Musau 6. E.
— Laurent, Gastwirth, Ludwigsgasse 24.
— Theobald, Tagner, Klebsauweg 3. E.
— Valentin, Wirth, R. Quartier 7. E.
Banhard Sebastian, Gypsfigurenmacher, s
Bankhauser Aug., Gärtner, Ziegelweg 29
Banschuh Anton, Handlungs-Commis, Fink
Banshaf Julius, Dienstmann, Stampfgasse
Bantelm Albert, Sattler, Artilleriewall
Banzet Ludw., Schreiner, Langestr. 19
Bapst Philipp, Specereihändler, Helleba
Barrack August, Prof. Dr., Oberbibliothek
Baraniki Aug., Postsecret., Allerheilige
Barasky Wilh., Eisenbahn-Sekretär, Köni
Barbaness Catharina, Rentnerin, Schla
— Georg, Gastwirth, Schleifhausplatz 33
— Louise, Rentnerin, Gewerkslaube 9.
Barbas geb. Heiser, Magd., Ww., Rentn
Barbier Bernhard, Octroi-Beamter, Lind
— Georg, Pfisterer, Bruderhofgasse 39.
— geb. Kientz, Magdalena, Ww., Tagn
— geb. Herrenberger, Maria Anna, Sch
— Nicolaus, Pflasterer, Blumenstaden 4.
Barchewitz Julie, Ww., Schildgasse 5.
Bardel Osc., Octr.-Einnehm., Schlacht
Bardes Ludwig, Metzger, Ballhausgasse 2
</column>
<column>
Bardol Anton, Schreiner, Finkweiler 3.
— Johann, Ackermann, Königshofen 34. E.
Bardotet Adolph, Maschinen-Techniker, f
Bareis Aug., Würstler, Schiffleutstader
Barmine geb. Barenne, Maria v., Ww., Re
Barfuss Friedr., Postschaffner, g. Renth
Bargholtz Wilhelm, Tagner, R. Quartier 2
Barkhausen Gustav, Kleiderhändler, Ver-

	A	B	C	D	E	F	G	H
	Nom	Prenom	Profession	Voie	Numero	annee	image	colonne
1	Balthes	Carl	Locomotivführer	Alterweimmarkt	10	1875	A_001875_0010	1
2	Baltzer	Carl	Professor	Krutenaustrasse	2	1875	A_001875_0010	1
3	Baltzinger	Daniel	Rentner	Langestrasse	93	1875	A_001875_0010	1
4	Baltzli	Michael	Wirth	Polygonstrasse	61	1875	A_001875_0010	1
5	Baltzli	Anna	Rentnerin	Magdalengasse	6	1875	A_001875_0010	1
6	Baltzli	Joh.	Tagner	Lingolsheimstr	19	1875	A_001875_0010	1
7	Baltzli	Georg	Brauknecht	Königshofen	44	1875	A_001875_0010	1
8	Baltzli	Georg	Taglöhner	Kinderspielg.	16	1875	A_001875_0010	1
9	Balucco	Heinr.	Hutmacher	Langestrasse	8	1875	A_001875_0010	1
10	Balzer	Georg	Bremser	Mittelhausbrgrst	2b	1875	A_001875_0010	1
11	Balzinger	Ludwig	Ackerer	Neuhof	93	1875	A_001875_0010	1
12	Balzli	Georg	Dienstmann	Thomasstadten	9	1875	A_001875_0010	1
13	Bambis	Dorethe	Wirthin	Johannesstaden	2	1875	A_001875_0010	1
14	Bambis	Emil	Schreiner	gr. Stadelgasse	6	1875	A_001875_0010	1
15	Bambis	Salome	Gärtnerin	gr. Rennegasse	12	1875	A_001875_0010	1
16	Bammes	Gottfried	Rentner	Neuer Kirchplatz	12	1875	A_001875_0010	1
17	Bandolzy	Carl	Postschaff.	Nikolausstad.	15	1875	A_001875_0010	1
18	Bangardt	Marie	Ww.	Eckboldsheimerstr	3	1875	A_001875_0010	1
19	Bangard	Michel	Schreiner- und G	Königshofen	51	1875	A_001875_0010	1
20	Bangard	Florenz	Tagner	Musau	6	1875	A_001875_0010	1
21	Bangratz	Laurent	Gastwirth	Ludwigsgasse	24	1875	A_001875_0010	1
22	Bangratz	Theobald	Tagner	Klebsauweg	3	1875	A_001875_0010	1
23	Bangratz	Valentin	Wirth	R. Quartier	7	1875	A_001875_0010	1
24	Banhard	Sebastian	Gypsfigurenmach	Spitalgasse	13	1875	A_001875_0010	1
25	Bankhauser	Aug.	Gärtner	Ziegelweg	29	1875	A_001875_0010	1
26	Banschuh	Anton	Handlungs-Comm	Finkmattstrasse	15	1875	A_001875_0010	1
27	Banshaf	Julius	Dienstmann	Stampfgasse	20	1875	A_001875_0010	1
28	Bantelm	Albert	Sattler	Artilleriewallstr	1	1875	A_001875_0010	1
29	Banzet	Ludw.	Schreiner	Langestr.	19	1875	A_001875_0010	1
30	Bapst	Philipp	Specereihändler	Hellebardengässc	2	1875	A_001875_0010	1
31	Barrack	August	Prof. Dr. Oberbib	Schlossplatz	2	1875	A_001875_0010	1
32	Baraniki	Aug.	Postsecret.	Allerheiligeng.	4	1875	A_001875_0010	1
33	Barasky	Wilh.	Eisenbahn-Sekret	Königshofen	43b	1875	A_001875_0010	1
34	Barbaness	Catharina	Rentnerin	Schlachthausplat	3	1875	A_001875_0010	1
35	Barbaness	Georg	Gastwirth	Schlachthausplat	3	1875	A_001875_0010	1
36	Barbaness	Louise	Rentnerin	Gewerkslaube	9	1875	A_001875_0010	1
37	Barbas	Magd.	Rentnerin	Gewerkslaube	23	1875	A_001875_0010	1
38	Barbier	Bernhard	Octroi-Beamter	Lindenfelsergasse	10	1875	A_001875_0010	1
39	Barbier	Georg	Pflasterer	Bruderhofgasse	39	1875	A_001875_0010	1
40	Barbier	Georg	Pflasterer	St. Nikolausgasse	18	1875	A_001875_0010	1
41	Barbier	Magdalena	Tagnerin	R. Quartier	317	1875	A_001875_0010	1
42	Barbier	Maria Anna	Schenkwirthin	Schiffelgasse	9	1875	A_001875_0010	1
43	Barbier	Nicolaus	Pflasterer	Blumenstaden	4	1875	A_001875_0010	1
44	Barchewitz	Julie	Ww.	Schildgasse	5	1875	A_001875_0010	1
45	Bardel	Osc.	Octr.-Einnehm.	Schlachthauspl.	2	1875	A_001875_0010	1
46								