



Un projet de recherche collaboratif entre

Historiens, économistes, démographes

Archivistes

Experts en intelligence artificielle



Et les services d'archives départementales et municipales

Christopher Kermorvant
Lionel Kesztenbaum

Financement sur **3 ans ½**



SocFace:

The local face of social change: one century of French social structure seen from the ground, 1836–1936

Collecter, traiter, retranscrire, organiser et analyser l'ensemble des listes nominatives du recensement de 1836 à 1936 (20 recensements).

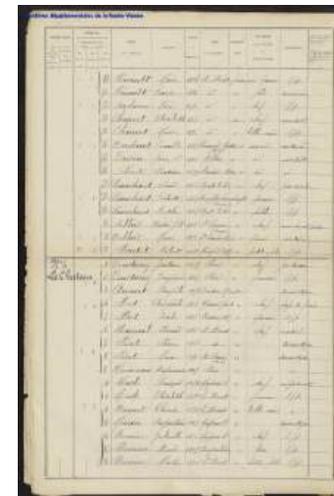
SocFace produira une base de données complète des individus ayant vécu en France entre 1836 et 1936 et l'utilisera pour analyser le changement social dans la longue durée.



A page from a 19th-century French census document, showing a grid of handwritten entries for individuals. The page is filled with dense, cursive handwriting, organized into columns and rows, typical of a population register.



A page from a 19th-century French census document, showing a grid of handwritten entries for individuals. The page is filled with dense, cursive handwriting, organized into columns and rows, typical of a population register.



A page from a 19th-century French census document, showing a grid of handwritten entries for individuals. The page is filled with dense, cursive handwriting, organized into columns and rows, typical of a population register.

De multiples objectifs et enjeux



Science

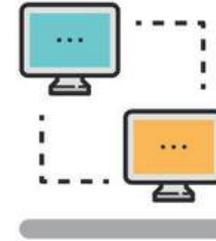
Etudes en économie, en histoire et en démographie : évolution du marché du travail, des inégalités, de la structure sociale, des mobilités.



Technologie

Reconnaissance automatique d'écriture manuscrite, analyse de tableau, traitement de plusieurs million d'images, accès à de multiples sources.

(TLR initial = 3)



Valorisation

Mise à disposition des données extraites en accès libre, versement aux propriétaires des fonds, valorisation par les portails.

DESIGNATION		NUMÉROS, PAR QUARTIER, VILLAGES, hameaux ou fermes.			NOMS	PRÉNOMS	ANNÉE	LIEU	NATIONALITÉ	SITUATION	PROFESSION.	REMARQUES
des quartiers, villages ou hameaux.	des fermes ou villages.	des numéros.	des messes.	des jardins.	DE FAMILLE.		de naissance.		LITTÉ.	PAR RAPPORT au chef de ménage.		
1	2	3	4	5	6	7	8	9	10	11	12	13
				2	Eribout	Lucie	1886	Reims	F	épouse	ans	
				3	Eribout	Senni	1909	Claris	"	enfant	"	
				4	Eribout	Alphonse	1912	"	"	"	"	
				5	Eribout	Reni	1916	Aubervilliers	"	"	"	
				6	Eribout	Audré	1920	"	"	"	"	
				7	Pierre	Henriette	1881	Reims	"	belle sœur	Magasinier	Boulonnais Valenciennes
				8	Pierre	Liliane	1912	Bagnollet	"	Nièce	ans	
				9	Pierre	Lucien	1918	Wizy-lez-Reims	"	neveu	ans	
				1	Parmentier	Justave	1877	Wafschateau	"	Chef	maçon	
				2	Parmentier	Lucie	1867	Lambach	"	épouse	ans	
				3	Parmentier	Maximilien	1885	Paris	"	enfant	Confectionnier	Chulman
				4	Parmentier	Emma	1900	W. Reims	"	"	empl.	Minist. Reims
				5	Parmentier	Ludovic	1902	"	"	"	plâtr.	Calix
				1	Caliez	Georges	1878	Lill	"	Chef	combustible	Magist.
				2	Caliez	Céline	1884	Tachy	"	épouse	ans	
				1	Renault	Jean	1869	Combray	"	Chef	chef équipe	Magist.
				2	Renault	Oliver	1875	Combray	"	"	"	

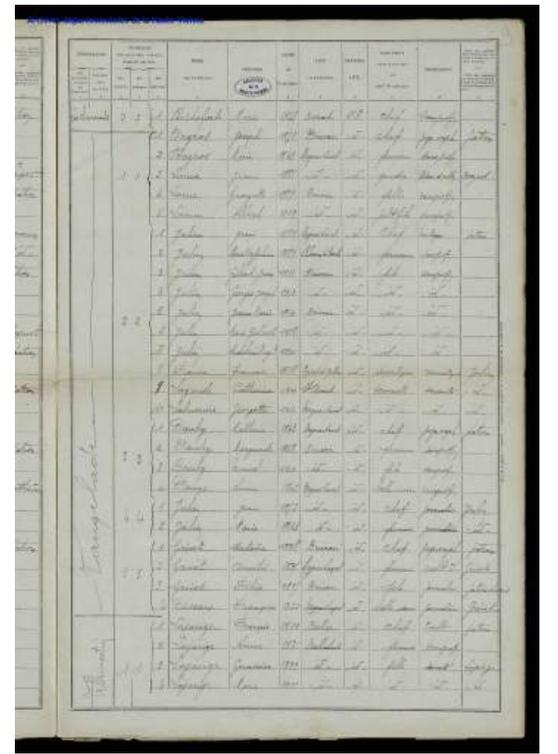
Une liste nominative:

Aubervilliers, 1921

Rue
des
Fillettes

Les listes nominatives: défis

- Documents décentralisés et nombreux scripteurs.
- Volume inconnu mais important (entre 5 et 15 M d'images).
- Conditions de conservation/numérisation variables.
- Lecture en 2 dimensions.
- Evaluation de l'erreur.
- Structurer, consolider, croiser l'information.
 - Utiliser la structure des listes et les informations agrégées pour contrôler la transcription.
 - Utiliser la stabilité temporelle pour améliorer les données.



The image shows a page from a handwritten ledger or list. The page is filled with rows of text, organized into columns. The handwriting is in a cursive script, typical of the 19th or early 20th century. The columns appear to contain names, dates, and other numerical or categorical data. The paper is aged and slightly yellowed. The text is written in dark ink, and the overall layout is that of a formal record-keeping document.

Lecture en 2 dimensions

312 — 4 —

DÉSIGNATION	NUMÉROS.				NOMS	ANNÉE	LIEU	NATIONAL.	SITUATION	PROFESSION.	Pour les pères, chefs d'entreprise, maîtres à domicile, instituteurs et autres.
	des carrés des villages ou hameaux	des rues des villages ou hameaux	des maisons	des familles							
Rue des Éclaireurs	11	11	2	Coribout	Lucie	1886	Reims	F	épouse	ans	
			3	Coribout	Léon	1909	Paris	"	enfant	"	
			4	Coribout	Alphonse	1912	"	"	"	"	
			5	Coribout	Reni	1916	subisillies	"	"	"	
			6	Coribout	Audie	1920	"	"	"	"	
			7	Pierre	Léonide	1881	Reims	"	bell. socie	Magasin	Parabonnie
			8	Pierre	Liliane	1918	Bagnollet	"	filie	ans	Alencien
			9	Pierre	Lucien	1918	Wizy-lez-Tours	"	revenu	ans	
			10	Parmenier	Justine	1877	Wizy-lez-Tours	"	chef	maçon	
			11	Parmenier	Lucie	1877	Lambach	"	épouse	ans	
			12	Parmenier	Maximilien	1885	Paris	"	enfant	Confectionner	Chulman
			13	Parmenier	Emma	1900	Wizy-lez-Tours	"	"	empl.	Minist. P. et M.
			14	Parmenier	Ludovic	1902	"	"	"	empl.	Calix
			15	Caliez	Jorges	1878	Lille	"	chef	Confectionner	Mag.
			16	Caliez	Clémence	1884	Éclaire	"	épouse	ans	
17	Renault	Jean	1869	Combray	"	chef	chef équipe	Mag.			

1. Rue
2. Maison
3. Ménage
4. Individu

Valorisation: retour aux archives (et au-delà...)

- Mise à disposition des données brutes aux archives
 - ❖ Sur une base nationale (sur FranceArchives).
 - ❖ Sur les bases des Archives Départementales.
 - ❖ Idéalement avec le lien vers les photos.
- Mise à disposition des données de la recherche
- Développement de nouvelles technologies pour la reconnaissance en masse
 - ❖ TLR final = 6/7.
- Ouverture sur d'autres sources: un modèle pour la mise à disposition des archives?

Merci pour votre attention

<https://socface.site.ined.fr/>

Christopher Kermorvant
kermorvant@tekliia.com

Lionel kesztenbaum
lionel.kesztenbaum@ined.fr