

Optimisation du traitement des notices d'autorité dans AtoM et IdRef



Mensch Mensch Mensch, Marion Bataillard, tempera sur bois, 2020-2021. (source : marionbataillard.com)

Retour d'expérience

Unil

UNIL | Université de Lausanne

Centre des littératures
en Suisse romande

Le CLSR en bref

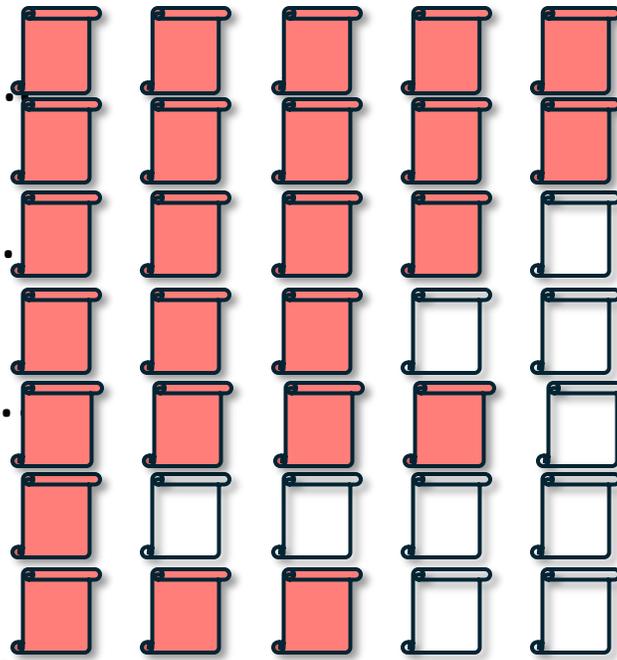
- Centre des littératures en Suisse romande de l'Université de Lausanne
- Fondé en 1965, anciennement Centre de recherches sur les Lettres romandes
- **Missions** : études littéraires romandes, travail sur archives, soutien à l'enseignement et la recherche
- Evolution vers la mission de conservation et de valorisation des fonds littéraires
- **Ressources humaines** : 1 directeur, 1 resp. de recherche, 1 chargée de recherche, 1 archiviste soutenu par 1 civiliste et, ponctuellement, 1 assistant·e étudiant·e

Le CLSR en bref

Fonds privés : écrivain·e·s, maisons d'édition, revues littéraires, acteur·rice.s du milieu littéraire romand

Types d'archives

- Manuscrits.....
-
- Correspondance.....
- .
- Photographies.....
- ..
- Documents audiovisuels.....
- Livres.....
-
- Objets.....



Etat des lieux

- Inventaires en 2019 : fichier Word (complet ou incomplet), tapuscrits, tableaux manuscrits, Filemaker

→ Pas d'inventaires en ligne

→ Données hétéroclites : structurées et non structurées

Fonds n° 226-2

1 page de <u>résumé</u>		p 226 → 248 La Maison Balagny Copie Rouvy		Une maison ancienne, une maison dans...	
p 226 → 233 Taby dans la Nuit Man.D. Version A		p 226 → 230 Maisons à l'eau Copie Devaion		chap. I	
p 1 → 13	p 1 → 14 VII ^e partie chap. I, Hologne	p 1 → 13	p 1-3 → 5	p 1 → 23	
p 14 → 23 Taby à l'aube		p 15 → 29 "La maison se réveille de bonne heure..." chap. II		p 15 → 29 p 14- chap II Le Lévrier	
p 242 → 248 [18 → 38] Je vous donnerai un nom nouveau "Comment les pensées de Taby..."		p 243 → 248 [24 → 29] chap. III		"Ils étaient cachés par leur nom mouai"	
p 25 → 27 [30 → 34] les deux silences T.O.M. "Une très soignée maison pour..."		p 231 → 234 [30 → 33]		p 29 → 42	
p 34 → 49 Maison à l'aube, La Maison		p 34 → 53		p 34 → 42 p 34-	

CORRESPONDANTS

CORRESPONDANTS ANSERMET, Ernest

LISTE

GROUPES Dossier GUENNE-CINGRIA Hélène

DATES 1956.12.20

COTES LR 1/1/131/1

Pour faire apparaître les index, cliquez sur les barres colorées

FONDS RESULTAT ANNULER QUITT

"La vie en cellule".
[1933 ?]. 3 f. ms. autogr., incomplet.

Deux textes sur le Corbusier. MSC 6

1. Texte probablement incomplet, numéroté de 26 A à 26 A14. Plusieurs chapitres d'un texte sur Le Corbusier. [1933?]. Ms. autogr. 18 f. de formats divers.

2. "Corbusier, précisions".
[1933?]. Ms. autogr. 6f. + notes diverses sur Le Corbusier. Ms. autogr. 7 f. non classées.
Tous ces textes peuvent éventuellement se rattacher à MSC 3.

"Pour parer au ravage des chemins de fer".
s.d. Ms. autogr., 1 cahier numéroté de 1 à 11.
[env. 1933-35 ?]

"Le Chrétien devant l'Art Moderne"
Ms. dactyl. avec adjonctions autogr. incomplet.
Le Ms. se compose de grandes "parties" divisées en chapitres.
Nous avons ici des traces des parties I, III, IV, VI, VII et ?
La plus complète est la III : la part du peuple.
[env. 1930].

Une partie de ces textes est parue dans NOVA ET VETERA:
I. Le goût du monde actuel, 1930, no 1, p. 38
II. La part du peuple, 1930, no 3, p. 277

MSC 9 +
91-917

Centre de recherches sur les lettres romandes
Fonds Gustave Rou

Suisse romande, 15 janvier 1938

« Récit »
« L'Hiver au village »
Ms 2 ff.
« Soirées au village », Lectures du foyer, janvier 1938

« Fernand contre le ciel »
« Introduction aux hymnes et poèmes de Hölderlin »
Dactyl. annoté 2 ff.
« Saisons, Fernand contre le ciel »
Photocopie 4 ff. (original fonds Daniel Simond)
Suisse romande, novembre 1937

« Sommeil d'Aimé »
« [La fête s'est dénouée comme une gerbe...] »
Ms 1 f.
Aujourd'hui, 27 mars 1930

« Pouvoirs d'une prairie »
« [...] danse lente et parfaite d'où tout écart...] »
Ms 1 f.
« [...] ve, bordée d'un rideau...] »
Dactyl. annoté 1 f.
« Guide du livre, octobre 1942

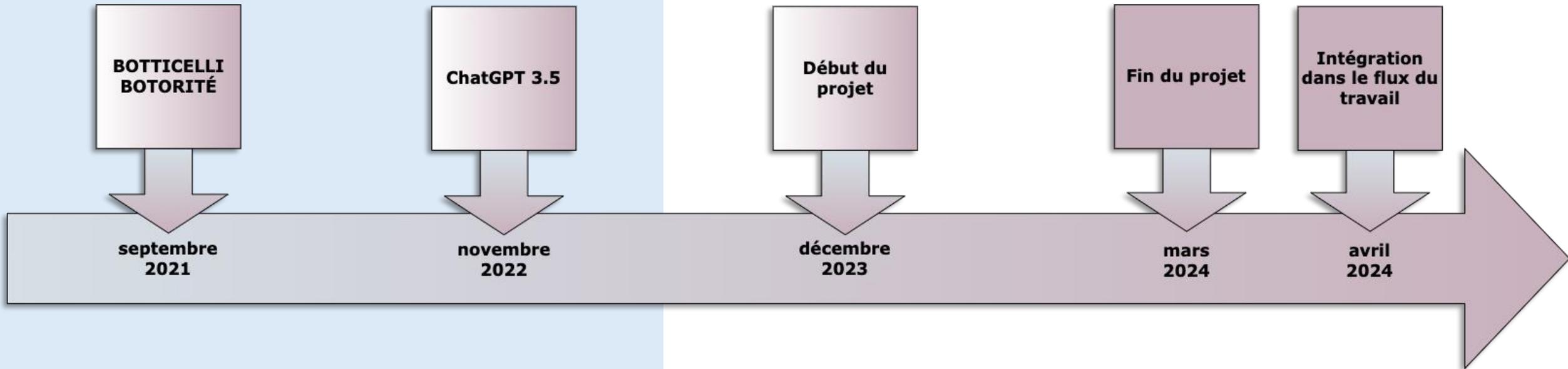
« D'une certaine pureté »
« [Ô fleuve te voici vêtu...] »
Ms 1 f. dans cahier, voir MS 4 C/14a
« Gustave Rouy : fragment d'une ébauche de "D'une certaine pureté" »
Ms 2 ff.

Mise en ligne des inventaires

- **PHŒBUS** sur AtoM
- Outil basé sur les normes archivistiques
- Soutien du Conseil international des Archives
- Développé par Artefactual
- Répond aux besoins de VALORISATION et de DIFFUSION



Chronologie



Botticelli/Botorité : indexation automatique de la correspondance

- Créateur : A. Mekacher
- But : indexer la correspondance des fonds (expéditeur et destinataire)
- Etapas clés : identification, tri, comparaison, indexation
- Uniquement la correspondance

ETAPE 1 : Acquisition des données

```
[3]: driver = Browser()  
importlib.reload(botticelli)  
botticelli.access_atom(driver)  
tagger = SequenceTagger.load('fr-ner-wikiner-0.4.pt')  
botticelli.browse_pages(driver, tagger, mode='Correspondance', includeList=['P014'])
```

2023-07-26 09:03:33,049 loading file fr-ner-wikiner-0.4.pt

ETAPE 2 : Filtrage des noms

```
[4]: importlib.reload(botticelli)  
entryDF = pd.read_csv('CLSR_Correspondance.csv')  
changeDict = {'Charles-Ferdinand Ramuz' : 'RAMUZ C.F.', 'Ramuz Charles-Ferdinand' : 'RAMUZ C.F.',  
botticelli.name_cleaning(entryDF, changeDict)
```

```
[5]: importlib.reload(botticelli)  
entryDF = pd.read_csv('CLSR_Correspondance.csv')  
botticelli.remove_duplicate_names(entryDF)
```

ETAPE 4 : Comparaison avec les notices d'autorité

```
[6]: importlib.reload(botticelli)  
driver = Browser()  
autoriteList = botticelli.autorite_retrieval(driver)
```

```
[7]: importlib.reload(botticelli)  
entryDF = pd.read_csv('CLSR_Correspondance.csv')#, sep=';'  
botticelli.autorite_comparison(entryDF, autoriteList)
```

```
[8]: importlib.reload(botticelli)  
entryDF = pd.read_csv('CLSR_Correspondance.csv')  
botticelli.find_previous_appearance(entryDF)
```

ETAPE 5 : Indexation automatique

```
[9]: driver = Browser()  
entryDF = pd.read_csv('CLSR_Correspondance.csv')  
  
importlib.reload(botticelli)  
botticelli.access_atom(driver)  
botticelli.indexation_automatique(driver, entryDF)
```

Série B - Correspondance

Trois conditions :

1. Classement des correspondances
 - Lettres rédigées par l'auteur
 - Lettres adressées à l'auteur
2. Intitulé (ISAD. 3.1.2) : NOM Prénom
3. Accès public du classement en XML

CH 000225-8 P014-VII-B - Correspondance

CH 000225-8 P014-VII-B-1 - Lettres rédigées par Charles Clément

CH 000225-8 P014-VII-B-1-BUD - BUDRY Paul

CH 000225-8 P014-VII-B-1-CLÉ1 - CLÉMENT Marie-Hélène

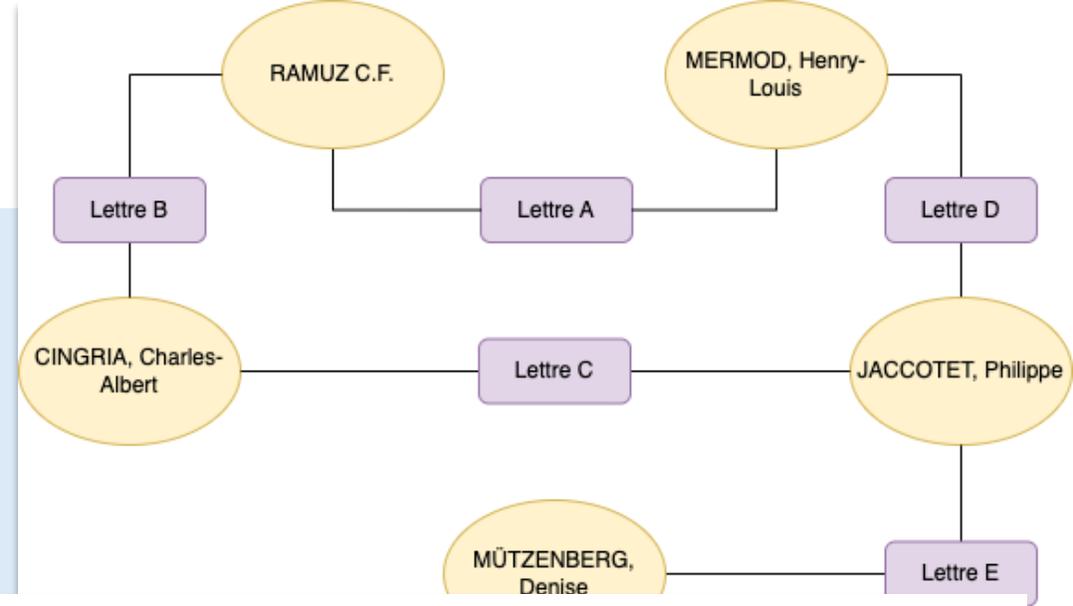
CH 000225-8 P014-VII-B-1-CLÉ2 - CLÉMENT Pierre-Paul

CH 000225-8 P014-VII-B-1-DUP - DUPLAIN Justin

	A	B	C	D	E
1	Titre	Cote	Profondeur	Autorites	Found
2	Correspondance	P014-VII-B	2		
3	Lettres rédigées par Charles Clément	P014-VII-B-1	3		
4	BUDRY Paul	P014-VII-B-1-BUD	4	["BUDRY Paul"]	[True]
5	CLÉMENT Marie-Hélène	P014-VII-B-1-CLÉ1	4	["CLÉMENT Marie-Hélène"]	[True]
6	CLÉMENT Pierre-Paul	P014-VII-B-1-CLÉ2	4	["CLÉMENT Pierre-Paul"]	[True]
7	DUPLAIN Justin	P014-VII-B-1-DUP	4	["DUPLAIN Justin"]	[True]
8	LONGCHAMP Marc	P014-VII-B-1-LON1	4	["LONGCHAMP Marc"]	[False]
9	LONGCHAMP Marguerite	P014-VII-B-1-LON2	4	["LONGCHAMP Marguerite"]	[True]
10	Années 1907-1916 : C. Clément à sa fiancée M. Longchamp	P014-VII-B-1-LON2-1	5		
11	Inventaire du deuxième classement	P014-VII-B-1-LON2-1-1	6		
12	Année 1907	P014-VII-B-1-LON2-1-2	6		
13	Année 1908	P014-VII-B-1-LON2-1-3	6		
14	Année 1910	P014-VII-B-1-LON2-1-4	6		
15	Année 1911	P014-VII-B-1-LON2-1-5	6		
16	Année 1912	P014-VII-B-1-LON2-1-6	6		
17	Année 1913	P014-VII-B-1-LON2-1-7	6		
18	Année 1914	P014-VII-B-1-LON2-1-8	6		
19	Année inconnue	P014-VII-B-1-LON2-1-9	6		
20	Année 1917	P014-VII-B-1-LON2-2	5		
21	Année 1922	P014-VII-B-1-LON2-3	5		
22	Années 1923-1924	P014-VII-B-1-LON2-4	5		
23	Année 1925	P014-VII-B-1-LON2-5	5		
24	Année 1926	P014-VII-B-1-LON2-6	5		
25	Année 1927	P014-VII-B-1-LON2-7	5		
26	Année 1928	P014-VII-B-1-LON2-8	5		
27	Année 1929	P014-VII-B-1-LON2-9	5		
28	Année 1930	P014-VII-B-1-LON2-10	5		
29	Année 1931	P014-VII-B-1-LON2-11	5		
30	Année 1932	P014-VII-B-1-LON2-12	5		
31	Années 1935-1936	P014-VII-B-1-LON2-13	5		
32	Enveloppes vides 1927-1936	P014-VII-B-1-LON2-14	5		
33	MORTON Harry	P014-VII-B-1-MOR	4	["MORTON Harry"]	[False]
34	Lettres adressées à Charles Clément	P014-VII-B-2	3		
35	AUBERJONNOIS René	P014-VII-B-2-AUB	4	["AUBERJONNOIS René"]	[True]
36	BOREL Paul	P014-VII-B-2-BOR	4	["BOREL Paul"]	[True]
37	CHINET Charles	P014-VII-B-2-CHI	4	["CHINET Charles"]	[True]
38	CLÉMENT Jeanne	P014-VII-B-2-CLE1	4	["CLÉMENT Jeanne"]	[False]
39	CLÉMENT Pierre-Paul	P014-VII-B-2-CLE2	4	["CLÉMENT Pierre-Paul"]	[True]
40	CLÉMENT-DUTOIT Emma	P014-VII-B-2-CLE3	4	["CLÉMENT-DUTOIT Emma"]	[False]
41	GAGNEBIN Elie	P014-VII-B-2-GAG	4	["GAGNEBIN Elie"]	[True]
42	GILLIARD Frédéric	P014-VII-B-2-GIL	4	["GILLIARD Frédéric"]	[True]
43	GOTTOFREY Victor	P014-VII-B-2-GOT	4	["GOTTOFREY Victor"]	[True]
44	HERMANJAT Abraham	P014-VII-B-2-HER	4	["HERMANJAT Abraham"]	[True]
45	LARBAUD Valery	P014-VII-B-2-LAR	4	["LARBAUD Valery"]	[True]
46	PONCET Marcel	P014-VII-B-2-PON	4	["PONCET Marcel"]	[True]
47	RAMUZ C.F.	P014-VII-B-2-RAM	4	["RAMUZ C.F."]	[True]
48	REYMOND Casimir	P014-VII-B-2-REY	4	["REYMOND Casimir"]	[True]

Résultat

- Lettre(s) liée(s) par deux notices d'autorité :
 - 1 notice liée au niveau du fonds
 - 1 notice liée au niveau de la pièce ou du dossier
- Nouvel accès aux fonds, par les notices d'autorité



Affiner les résultats par :

FAIT PARTIE DE	
Tout	
Mermod, Henry-Louis	8
Buchet, Gérard	3
Formes et Couleurs (André Held)	1
Jaccottet, Philippe	1
Jotterand, Franck	1
Robert, Steven-Paul	1
Auberjonois, René	1
Cornuz, Jeanlouis	1
Lachenal, François	1

NOM	
Tout	
ROUD Gustave	18
ROY Bruno	1
CORNUZ Jeanlouis	1
ROCHAT Jean	1
VAN MUYDEN Emilie	1
CLERC Jean-Louis	1

NIVEAU DE DESCRIPTION	
Tout	
Dossier	10
Pièce	5

Affichage de 18 résultats

Fonds d'archives

ROUD Gustave x

Options de recherche avancée

[Aperçu avant impression](#) [Hierarchy](#) [Exporter au format CSV](#) Affichage :  

Trier par : [Intitulé](#) Dire

Bruno ROY à Gustave ROUD

CH 000225-8 P023-1-2-1 · Pièce · 1976 ·

Fait partie de [Buchet, Gérard](#)

1 carte de vœux des Éditions Fata Morgana, avec une linogravure de Raoul Ubac.

Carte collective à Gustave ROUD

CH 000225-8 P023-1-2-2 · Pièce · 1942 ·

Fait partie de [Buchet, Gérard](#)

1 carte collective adressée signée par Daniel SIMOND, Pierre EMMANUEL, Edmond GILLIARD, Jean DESCOULLA MARTIN, P. BURNIER (?), Gérard BUCHET, Constantin MAVROMICHALIS, ...)

Correspondance de Gustave Roud à Jeanlouis Cornuz

CH 000225-8 P096-C-2-4 · Dossier · 1950-1976 ·

Fait partie de [Cornuz, Jeanlouis](#)

Diverses lettres ms.

MERMOD Henry-Louis à ROUD Gustave

CH 000225-8 P059-B-4-LOR-5 · SDossier · 1949-03 ·

Fait partie de [Mermod, Henry-Louis](#)

Lieu(x) : Lausanne ; Collation : 2 l.d.s ;

Ci-joint : Couverture de la chemise d'origine du dossier, mention a. de HLM : " Lorelei - anthologie de poésie allemande ".

CINGRIA Amélie cingria_amelie · Personne	
CINGRIA Delphine cingria_delphine · Personne	
CLAVÉ Tony C. de clave_tony_c_de · Personne	
DE HOOP SCHEFFER de_hoop_scheffer · Personne	
DELAPORTE Alice delaporte_alice · Personne	
DELIRY Paul deliry_paul · Personne	
DELMOULY Marcel	
DELMOULY Renée	
DEREY Léon	
DEROFF Emile deroff_emile · Personne Aurait été un proche de l'Union des Armes en 1916.	
DUC Jacqueline duc_jacqueline · Personne	
ECUVILLON Michel Ecuillon_Michel · Personne	
EIMANN Jean eimann_jean · Personne	
EIMANN Julia eimann_julia · Personne	
EIMANN Marguerite eimann_marguerite · Personne	

Bilan intermédiaire

- Inconvénients/Problèmes :
 - Risque accru des doublons
 - Indexation de personnes inconnues, avec peu d'information
 - Notices d'autorité vierges

SOLUTIONS	OUTILS	DEFAUTS
nettoyage/enrichissement manuel	OpenRefine LibreOffice recherche internet	Chronophage Exténuant
enrichissement automatique (<i>scrapping</i>)	OpenRefine LibreOffice Python IdRef	Manque de compétence informatique

Enrichissement automatique

- Coder en Python à l'aide de ChatGPT 3.5
- IdRef = identifiants et référentiels pour l'enseignement supérieur et la recherche (ESR)

→ *Base de données mettant à disposition le contenu de milliers de notices d'autorité et une palette d'outils pour y accéder.*

The screenshot displays the IdRef search interface. At the top left is the IdRef logo with the text 'Identifiants et Référentiels pour l'enseignement supérieur et la recherche'. To the right are icons for 'paprika.idref.fr', 'data.idref.fr', and 'Documentation', along with input fields for 'IDENTIFIANT' and 'MOT DE PASSE'. Below this is a navigation bar with 'Nouvelle recherche', 'Modifier la recherche', and 'Créer une notice'. The main area is divided into three sections: 'Type d'autorité' with radio buttons for categories like 'Nom de personne', 'Nom de collectivité', etc.; 'Termes de recherche' with a search box and a 'Recherche' button; and 'Filtres' with input fields for 'Langue' and 'Pays', and radio buttons for 'En vie' (Oui, Non, Indifférent).

Buts

- Compléter les notices d'autorité de Phœbus
- Lier Phœbus à IdRef
- Lier IdRef à Phœbus

Zone du contrôle

Identifiant de notice d'autorité	ebner-jeannie_idref-180716514
Entretenu par	Centre des littératures en Suisse romande
Statut	Final
Dates de production, de révision et de suppression	Notice Phœbus. Dernière mise à jour : 2024-05-15 à 10:53:45
Sources	Voir également l'Agence bibliographique de l'Enseignement supérieur (ABES) : http://www.idref.fr/180716514/id .

Données et archives scientifiques



Archives littéraires du CLSR-UNIL

1 référence PHŒBUS : Tout voir dans PHŒBUS

- EBNER Jeannie

Etapes clés

- 1) prendre la liste des autorités à compléter
- 2) faire des requêtes SPARQL dans IdRef
- 3) récupérer les résultats des requêtes
- 4) **trier les notices : un, plusieurs résultats et aucun résultat**
- 5) vérifier les « alignements »* des autorités
- 6) saisir ces informations dans Phœbus

*alignement : correspondance entre une notice d'autorité sur Phœbus et une notice d'autorité IdRef

Phase de tâtonnement

- **SKYSCRAPER** : Un bot qui va récupérer les informations sur IdRef
- **BOTORITE-2.0** : Un bot pour saisir les informations récupérées sur Phœbus
- **BMMW** (Bot Multi Matches Work) : Un bot pour traiter les résultats multiples
- **SAW** (Scrap And Write) : Un bot pour lier IdRef à Phœbus



Le processus se complexifie

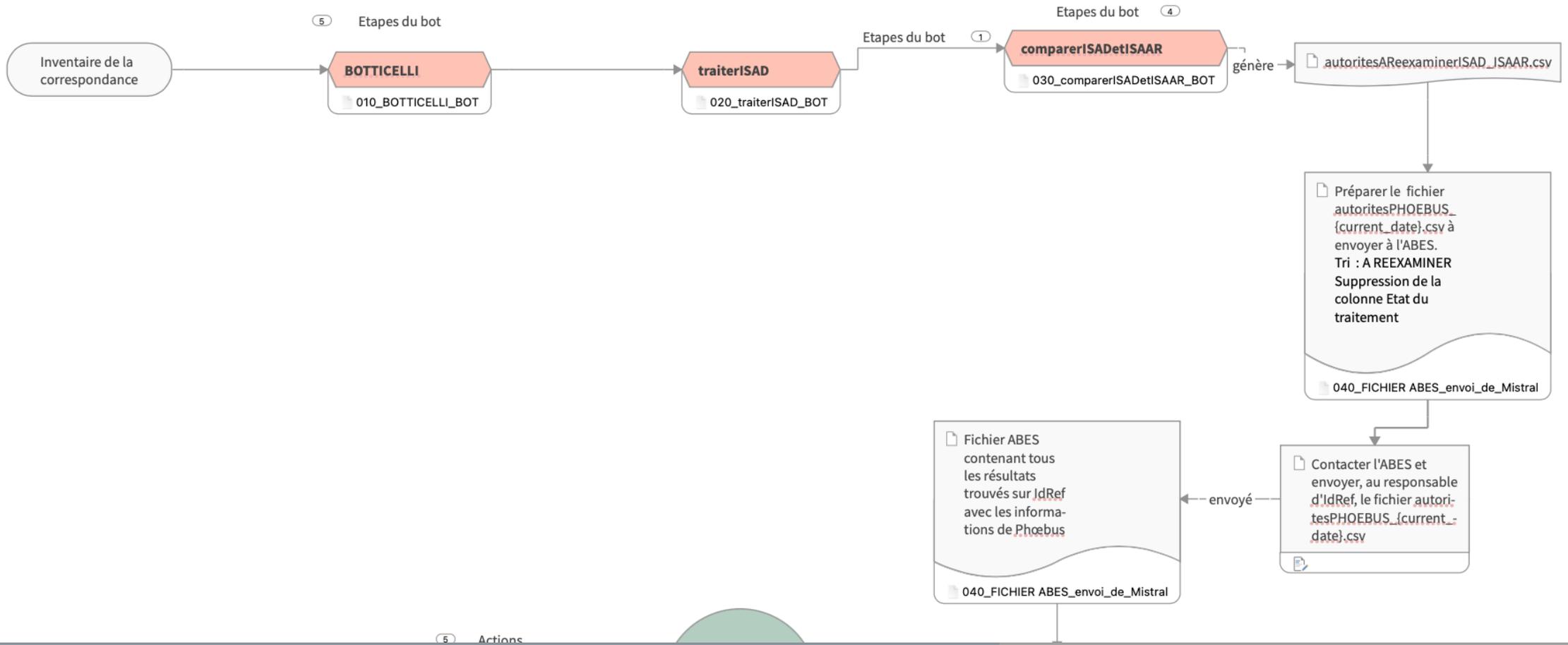
Etapes clés

- 1) prendre la liste des autorités à compléter 
- 2) faire des requêtes SPARQL dans IdRef 
- 3) récupérer les résultats de requêtes 
- 4) **trier les notices : un, plusieurs résultats et aucun résultat** 
- **5) vérifier les « alignements » des autorités**
→ quid des homonymes ? 
- 6) saisir ces informations dans Phœbus 

La bonne approche

- « Quelles sont les informations dont j'ai besoin pour confirmer un alignement ? »
- Changement de logique : traitement des autorités à compléter [ISAAR-CPF] → traitement de toutes les lettres indexées [ISAD-G]
- Contacter le responsable IdRef-autorité : F. Mistral
- CLSR : fournit une liste de toutes les descriptions de lettres qui sont indexées, avec les informations utiles
- IdRef : retourne une liste de tous les résultats trouvés sur IdRef, avec les informations utiles

Workflow : de l'inventaire des correspondances à l'envoi à IdRef



Liste fournie par IdRef

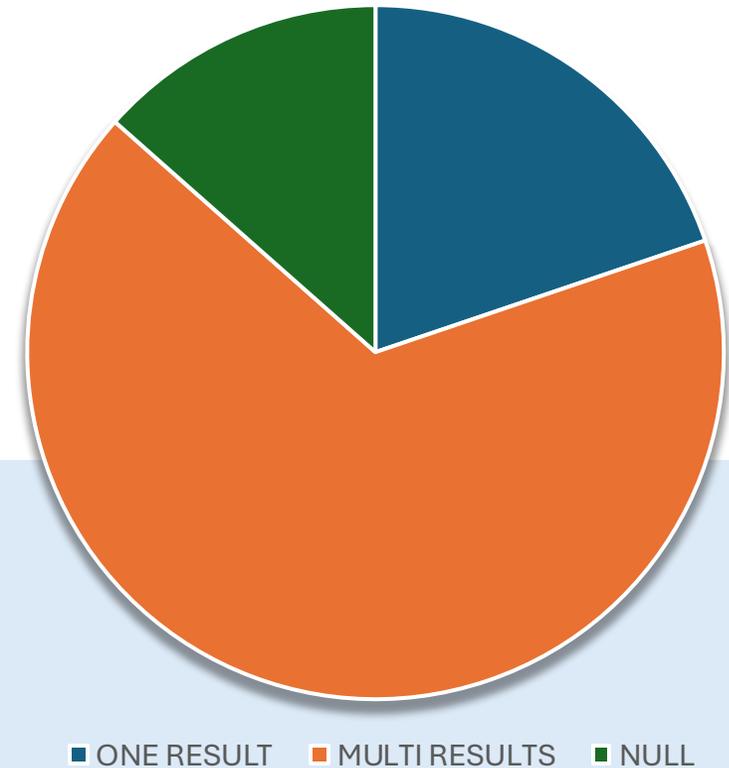
Fichier contenant tous les résultats trouvés sur IdRef avec les informations de Phœbus

File	pour tri	nameAccessPoints	title	période	ID_IdRef	Appellat	Langue	Pays	Naissanc	Déc	Note300	Note340	Source	Type	scopeAn	eventSt	eventEn	eventDa	referenceCode	requete	Idref	
clsr-0001_2	PEYRON Gisèle	PEYRON Gisèle	PEYRON Gisèle	1960-1962	132633965	Peyron, Gisèle	fre	CH	19XX	19XX			Membre fondateur de l'A	Bibliographie de Chai	2 lettres ma	1960	1962	1960-1962	P059-G-2-b-PEY		https://www.idref.fr/Sru	
clsr-0001_2	ZERMATTEN Maurice	ZERMATTEN Maurice	ZERMATTEN Maurice	1955	027203603	Zermatten, Maurice	fre	CH	19101002	200102			Romancier, essayiste, aut	Vue panoramique sur a	1 lettre mar	1955	1955	1955	P059-G-2-b-ZER		https://www.idref.fr/Sru	
clsr-0001_2	ANCEL Suzanne	ANCEL Suzanne	ANCEL Suzanne	1933-1956	245426752	ANCEL Suzanne	fre	FR					Titulaire d'un doctorat	Recherche sur la toxita	20 lettres e	1933	1956	1933-1956	P063-B-1-ANC		https://www.idref.fr/Sru	
clsr-0001_2	GILLIARD Edmond	GILLIARD Edmond	GILLIARD Edmond	1929-1936	02689162x	Gilliard, Edm	fre	CH	1875	1969			Enseignant, écrivain, essa	La Dramatique du mc a	2 cartes pos	1929	1936	1929-1936	P063-B-1-GIL		https://www.idref.fr/Sru	
clsr-0001_2	MERMOUD Albert	MERMOUD Albert	MERMOUD Albert	1959	02933456x	Mermoud, Albert	fre	CH	1905				Éditeur suisse	BN APP (en ligne), 19	1 lettre à Al	1959	1959	1959	P063-B-1-MER		https://www.idref.fr/Sru	
clsr-0001_2	SCHUDEL Otto	SCHUDEL Otto	SCHUDEL Otto	1955	06140850E	Mermoud, Albert	fre	FR	1931					Mémoire du Mont-Bli	1 lettre à Ot	1955	1955	15 mai 195	P063-B-1-SCH		https://www.idref.fr/Sru	
clsr-0001_2	ABRAHAM Pierre	ABRAHAM Pierre	ABRAHAM Pierre	1952	114911932	Abraham, Pierre	fre	FR	19XX					Manuel prémedical p a	1 lettre.	1952	1952	4 novembre	P063-B-2-ABR		https://www.idref.fr/Sru	
clsr-0001_2	ABRAHAM Pierre	ABRAHAM Pierre	ABRAHAM Pierre		12298324	Abraham, Pierre	fre	FR	18910311	196601			Architecte en chef des b	Pol Abraham, Centre								
clsr-0001_2	ABRAHAM Pierre	ABRAHAM Pierre	ABRAHAM Pierre		02917356E	Abraham, Pierre	fre	FR	1936	200307			Auteur, responsable de	Armen / J.-P. Abraham								
clsr-0001_2	ABRAHAM Pierre	ABRAHAM Pierre	ABRAHAM Pierre		074001973	Abraham, Pierre	fre	FR	19610915				Praticien hospitalier au s	Hémodynamique veii								
clsr-0001_2	ABRAHAM Pierre	ABRAHAM Pierre	ABRAHAM Pierre		026647834	Abraham, Pierre	fre	FR	18920301	197405			Homme de lettres. A été c	Figures : recherches s								
clsr-0001_2	ABRAHAM Pierre	ABRAHAM Pierre	ABRAHAM Pierre		10878634x	Jonain, Pierre	fre	FR	1799	1884			Linguiste	Dictionnaire du patoi								
clsr-0001_2	ABRAHAM Pierre	ABRAHAM Pierre	ABRAHAM Pierre		02815569E	Poterat, Albert	fre	FR	17	18			Capitaine de vaisseau	Traité pratique à l'usa								
clsr-0001_2	ABRAHAM Pierre	ABRAHAM Pierre	ABRAHAM Pierre		18512635E	Pury, Abraham	fre	CH	17240703	180702			Officier au service de Sarc	Dict. historique de la								
clsr-0001_2	ABRAHAM Pierre	ABRAHAM Pierre	ABRAHAM Pierre		223376604	Pierre de la	lat	NL	16100315	168310			Carme déchaussé	Les fleurs du Carmel, i								
clsr-0001_2	ABRAHAM Pierre	ABRAHAM Pierre	ABRAHAM Pierre		09241972C	Pajon de M	fre	FR	17..	17..			Médecin de l'Hôtel de	Lieu de naissance Blois								
clsr-0001_2	ANCEL Suzanne	ANCEL Suzanne	ANCEL Suzanne	1934-1961	031306853	Wolfgang, F	dut	NL	163403	169401			Imprimeur-libraire holla	Ledeboer. Marques P								
clsr-0001_2	ANCEL Suzanne	ANCEL Suzanne	ANCEL Suzanne	1941-1942	245426752	ANCEL Suzanne	fre	FR					Titulaire d'un doctorat	Recherche sur la toxita	19 lettres d	1934	1961	1934-1961	P063-B-2-ANC		https://www.idref.fr/Sru	
clsr-0001_2	BALLARD Jean	BALLARD Jean	BALLARD Jean (Cahiers du Sud)	1941-1942	185133785	Ballard, Jean	fre	XX					Auteur d'une thèse de do	Approche géologique a	4 lettres de	1941	1942	1941-1942	P063-B-2-BAL		https://www.idref.fr/Sru	
clsr-0001_2	BALLARD Jean	BALLARD Jean	BALLARD Jean		24379950C	Ballard, Jean	fre	FR					Titulaire d'un mémoire et	L'insaisissabilité du s								
clsr-0001_2	BALLARD Jean	BALLARD Jean	BALLARD Jean		02670256E	Ballard, Jean	fre	FR	18931114	19730E			Fondateur et directeur de	Correspondance 193								
clsr-0001_2	BALLARD Jean	BALLARD Jean	BALLARD Jean		13240379x	Ballard, Jean	fre	FR	1780	1842			Médecin ordinaire de la g	Principes de médecin								
clsr-0001_2	BALLARD Jean	BALLARD Jean	BALLARD Jean		035816953	Ballard, Chr	fre	FR	1772	182512			Imprimeur-libraire ; impi	Delalain, Paris, 1789 -								
clsr-0001_2	BALLARD Jean	BALLARD Jean	BALLARD Jean		031108407	Ballard, Chr	fre	FR	1701	176505			Autre(s) graphie(s) : Ball	Lotting. Lepreux, Paris								
clsr-0001_2	BALLARD Jean	BALLARD Jean	BALLARD Jean		19781646C	Bauché, Thé	fre	FR	1760 ?	1829			Imprimeur-libraire. Nom	De veritate / Charles -								
clsr-0001_2	BALLARD Jean	BALLARD Jean	BALLARD Jean		165174072	Ballard, Jean	fre	FR	193X	2022			Acteur	Les contes du march								
clsr-0001_2	BALLARD Jean	BALLARD Jean	BALLARD Jean		031108423	Ballard, Jean	und	FR	1674 ?	175005			Imprimeur-libraire ; (seu	La clef des chansonni								
clsr-0001_2	BALLARD Jean	BALLARD Jean	BALLARD Jean		06733025E	Ballard, Ve	und	FR	17..	179. ?			Nom de jeune fille : Mari	Lotting. Almanach de l								
clsr-0001_2	BEAUSIRE Pierre	BEAUSIRE Pierre	BEAUSIRE Pierre	1946	02671292x	Beausire, Pi	fre	CH	1902	1990			Poète et critique littér	Mallarmé : poésie et	2 lettres.	1946	1946	1946	P063-B-2-BEA		https://www.idref.fr/Sru	
clsr-0001_2	BILLIER Jacques	BILLIER Jacques	BILLIER Jacques	1951			fre	FR							1 lettre.	1951	1951	Août 1951	P063-B-2-BIL		https://www.idref.fr/Sru	
clsr-0001_2	BOHNENBLUST Gottfried	BOHNENBLUST Gottfried	BOHNENBLUST Gottfried	1932-1953	060775092	Bohnenblust	ger	CH	1883	1960			Écrivain suisse, égaleme	Deutsche Biographisc	15 lettres, c	1932	1953	1932-1953	P063-B-2-BOH1		https://www.idref.fr/Sru	
clsr-0001_2	BOHNENBLUST Theodor	BOHNENBLUST Theodor	[BOHNENBLUST] Theodor	1931-1934			fre	FR							2 lettres.	1931	1934	1931-1934	P063-B-2-BOH2		https://www.idref.fr/Sru	
clsr-0001_2	BONNARD André	BONNARD André	BONNARD André	1949	19711377x	Bonnard, Andr	fre	FR	19420119				Médecin	Problèmes médicaux	1 carte de b	1949	1949	1949	P063-B-2-BON1		https://www.idref.fr/Sru	
clsr-0001_2	BONNARD André	BONNARD André	BONNARD André		02673785x	Bonnard, Andr	fre	CH	18880816	196012			Helléniste. - Écrivain et	Civilisation grecque /								
clsr-0001_2	BONNARD Georges	BONNARD Georges	BONNARD Georges	1929	07489314E	Bonnard, Ge	fre	FR	1886				Angliciste.	Le journal de Gibbon	1 lettre de	1929	1929	23 octobre	P063-B-2-BON2		https://www.idref.fr/Sru	
clsr-0001_2	BONNARD Georges	BONNARD Georges	BONNARD Georges		24224147E	Bonnard, Ge	fre	FR	19220415				Titulaire d'une thèse en	De l'opération césari								
clsr-0001_2	BOVET Lucien	BOVET Lucien	BOVET Lucien	1931-1950	081607334	Bovet, Lucie	fre	CH	1962XXXX				Biologiste et musicien	Aide-mémoire de bio	5 lettres et	1931	1950	1931-1950	P063-B-2-BOV		https://www.idref.fr/Sru	
clsr-0001_2	BOVET Lucien	BOVET Lucien	BOVET Lucien		263673197	Bovet, Jacq	fre	CH	1934XXXX				Ne pas confondre avec	Biologiste. Docteur de l'								
clsr-0001_2	BOVET Lucien	BOVET Lucien	BOVET Lucien		200471821	Bovet, Lucie	fre	CH	19070920	195107			Psychiatre, spécialement	Les enfants nerveux :								
clsr-0001_2	BOVET Lucien	BOVET Lucien	BOVET Lucien		252878922	Bovet, Jacq	fre	CH	19XX				A écrit aussi en allem	Zoologue. Docteur de l'								
clsr-0001_2	CABANIÉ Urbain	CABANIÉ Urbain	CABANIÉ Urbain	1943			fre	FR							1 lettre sign	1943	1943	20 novemb	P063-B-2-CAB		https://www.idref.fr/Sru	
clsr-0001_2	CAMPICHE Pierre	CAMPICHE Pierre	CAMPICHE Pierre	1949-1955	27425683E	Campiche, I	fre	CH					Entrée illégale et séjo	Dossier con	1949	1955	1949-1955	P063-B-2-CAM		https://www.idref.fr/Sru		
clsr-0001_2	CLERC Charly	CLERC Charly	CLERC Charly	1937	026791501	Clerc, Charl	fre	CH	18820815	195810			A aussi traduit du latin	Poète, romancier, auteur	1 lettre.	1937	1937	26 juin 193	P063-B-2-CLE		https://www.idref.fr/Sru	
clsr-0001_2	COTTON Eugénie	COTTON Eugénie	COTTON Eugénie	1951	077904125	Cotton, Eug	fre	FR	1881	1967			physicienne, pédagogue,	http://www.int-evry.	1 lettre du	1951	1951	1951	P063-B-2-COT		https://www.idref.fr/Sru	
clsr-0001_2	DESCOULLAYES Jean	DESCOULLAYES Jean	DESCOULLAYES Jean	1933-1961	059660961	Descoullay	fre	CH	19030910	19610E			Critique théâtral et critiq	Oeuvres complètes / l	6 lettres.	1933	1961	1933-1961	P063-B-2-DES		https://www.idref.fr/Sru	
clsr-0001_2	ECKARDT Léonore	ECKARDT Léonore	ECKARDT Léonore	1948			fre	FR							1 lettre à la	1948	1948	31 mars 19	P063-B-2-ECK		https://www.idref.fr/Sru	
clsr-0001_2	ERNST Fritz	ERNST Fritz	ERNST Fritz	1940	195822471	Winkelman	ger	DE	19011102				Médecin	Zur Pathogenese und	1 lettre.	1940	1940	15 mars 19	P063-B-2-ERN		https://www.idref.fr/Sru	
clsr-0001_2	ERNST Fritz	ERNST Fritz	ERNST Fritz		05033591x	Ernst, Fritz	ger	DE	1905	1963			Historien allemand	BN APP (en ligne), 19								
clsr-0001_2	ERNST Fritz	ERNST Fritz	ERNST Fritz		035554363	Ernst, Fritz	ger	DE	18890614	195803			Professeur de littérature	Forschungsprobleme								
clsr-0001_2	ERNST Fritz	ERNST Fritz	ERNST Fritz		27088444C	Noether, Fri	ger	DE	1884	1941			Mathématicien allemand	Über die Theorie des l								
clsr-0001_2	ERNST Fritz	ERNST Fritz	ERNST Fritz		067022847	Schmid, Err	ger	DE	19040307	196001			Musicologue et chef d'ori	Sonaten Klavier und V								
clsr-0001_2	ERNST Fritz	ERNST Fritz	ERNST Fritz		251535657	Fürbringer,	ger	DE	19000727	198810			Acteur	Die Gruft mit dem R								

Statistique

Données envoyées à IdRef	4 318
1 donnée CLSR → 1 résultat	1 976
1 donnée CLSR → plusieurs résultats	6 668
1 donnée CLSR → aucun résultat	1 342
Totale de données IdRef	<u>9 986</u>

Données récupérées [output]



Problème vertigineux

La vérification
prendra une éternité
et épuisera tout le
personnel

Francesco Cairo, *Agony in the Garden*, huile sur toile, circa 1630-1667



Solutions



Jean-François Millet, *Les glaneuses*, 1857, huile sur toile. (source : wikiart.org)

- Faciliter au maximum la vérification des alignements
- Ajouter des colonnes pour trier les données
- Fournir des probabilités de *match*
- Éliminer les *mismatches*
- Rendre le tableau Excel *User friendly*
- Travailler en équipe

Pour le tri

- Comparaison entre la **date de naissance** / la **date de décès** et la **date de production** de la lettre
- Âge par rapport à la production de la lettre (< ou > 15 ans)
- Date de naissance **avant** ou **après 1850**
- Nombre de lettres décrites dans Phoebus

Probabilité

- Attribution d'un pourcentage augmentant la probabilité d'un *match*

Critères/ Probabilité (%)	NOMS	STATUT	Date de N. / prod. Lettre	Date de D. / prod. Lettre	Age par rapport à la prod. de la lettre	Nbre de lettre dans Phoebus	Contexte du fonds (expéd ou destin.)	Pays ?	Naissance avant 1850 ?	PRO- BABI- LITÉ
Attribution	+50 %	ONE RESULT, MULTI RESULTS, NULL	Lettre après naissance = +5 % Lettre avant naissance = -45 % (anomalie) Pas de date de naissance = 0% (info insuffisante)	Lettre avant décès = +5 % Lettre après décès = -45 % (anomalie) Pas de date de décès = 0% (info insuffisante)	+ 15 ans = +5 % ; - 15 ans = -5 %	une occurrence = +5% plusieurs = +10%	+10%	CH = +5 % FR = +2 %	NON = + 5% OUI = -50 %	95% max

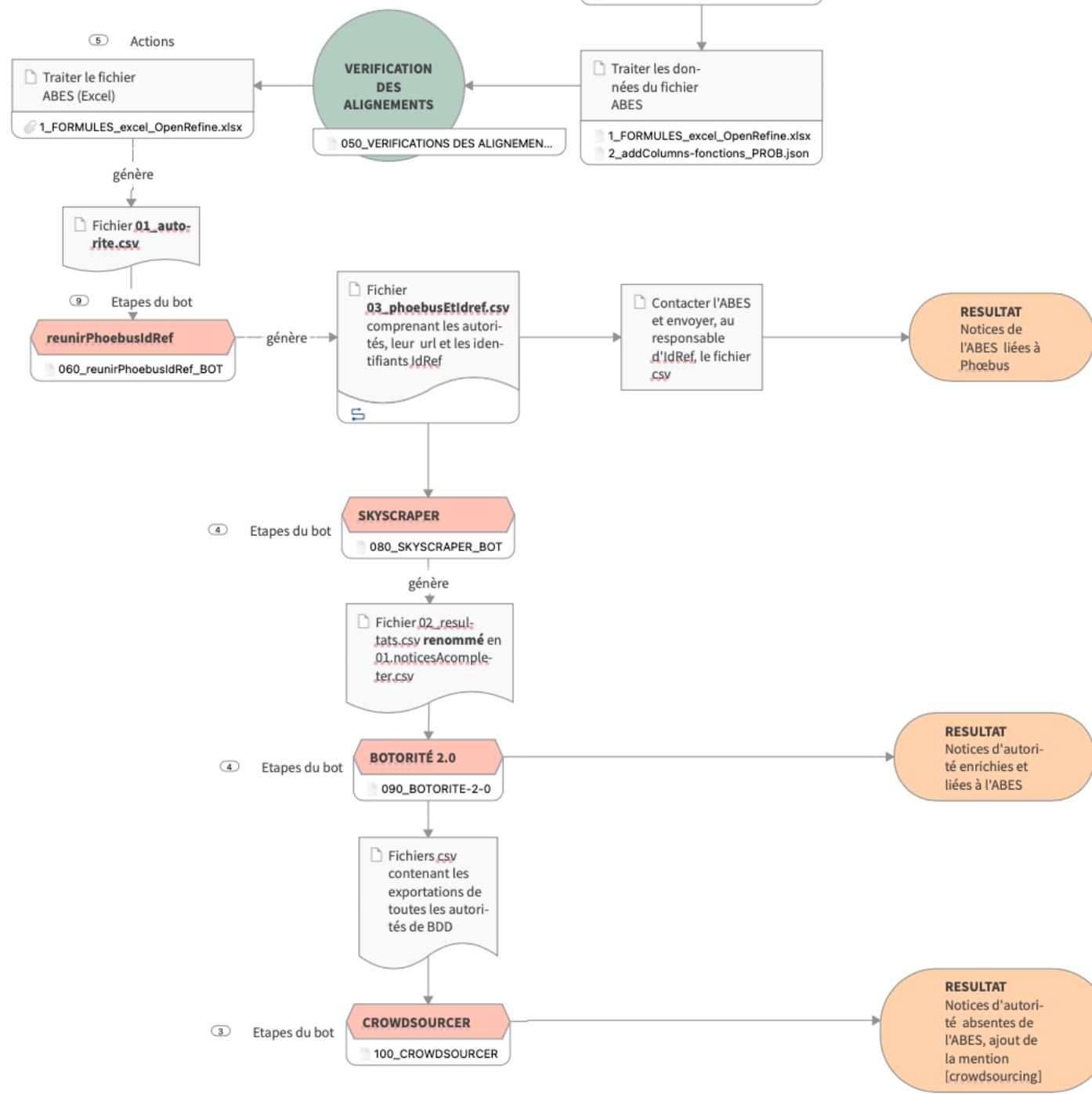
Le fichier Excel «*user friendly*»

A	B	C	D	E	M	N	O	P	Q	R	S	T	U	V	W
nameAccessPoints	pour tri	ALIGNEMENT	ID_IdRef	Prob	Appellation_IdRef	angu	Pays	Naissance	Décès	Note300	Note340	Source	ISAD_title	scopeAr	
353 IANOZ Michel	Ianoz Michel	MISMATCH	1 idref-050373285	65%	Ianoz, Michel V (1936-....)	eng	XX	1936				Spécialiste de la compatibilité électro EMC analysis methods ar	IANOZ Michel	2 l.	
354 IFFLAND Hugo	Iffland Hugo	MISMATCH	0 idref-	0%									IFFLAND Hugo	2 l.	
355 IMBS Françoise	Imbs Françoise	MISMATCH	0 idref-	0%									IMBS Françoise	1 l.	
356 IMER André	Imer André	MISMATCH	0 idref-	0%									IMER André	13 lettres, p	
357 INEICHEN Hugo	Ineichen Hugo	MISMATCH	0 idref-	0%									INEICHEN Hugo	1 lettre, de	
358 ISOZ Pierre	Isoz Pierre	MISMATCH	2 idref-273883399	65%	Isocz, Pierre	fre	CH					Quelques pas dans l'histc ISOZ Pierre	ISOZ Pierre	1 lettre.	
359	Isoz Pierre	MISMATCH	idref-242197124	62%	Isocz, Frédéric-Pierre	fre	FR				Psychanalyste	Pourquoi croyons-nous ? : Dialogue en liberté d'un philosophe et d'un			
360 ISSAC Jean-Jacques	Issac Jean-jacques	MISMATCH	0 idref-	0%									ISSAC Jean-Jacques	1 l.	
361 IZZO Carlo	Izzo Carlo	MISMATCH	2 idref-243708521	60%	Izzo Puebla,										
362	Izzo Carlo	MATCH	idref-195378423	70%	Izzo, Carlo (
363 JABON Laure	Jabon Laure	MISMATCH	0 idref-	0%											
364 JACCARD Charles-Henri	Jaccard Charles-hen	MISMATCH	0 idref-	0%											
365 JACCARD Claude	Jaccard Claude	MISMATCH	3 idref-14635589X	65%	Jaccard, Cla										
366	Jaccard Claude	MISMATCH	idref-255971001	65%	Jaccard, Cla										
367	Jaccard Claude	MATCH	idref-272700665	70%	Jaccard, Je										
368 JACCARD Marc	Jaccard Marc	MISMATCH	1 idref-191785016	60%	Jaccard, Marc										
369 JACCARD Olivier-Marc	Jaccard Olivier-marc	MISMATCH	0 idref-	0%											
370 JACCARD Roland	Jaccard Roland	MATCH	1 idref-026932369	72%	Jaccard, Roland										
371 JACCARD-VILLARD Juliette	Jaccard-villard Juliet	MISMATCH	1 idref-164150854	22%	Chevillard, Juliette (1942-1971)	fre	FR	19421130	19710316			Actrice	http://fr.wikipedia.org/wiki/JACCARD-VILLARD Juliette	1 l.	
372 JACCOTTET Catherine	Jaccottet Catherine	MISMATCH	1 idref-031679129	65%	Jaccottet, Catherine (19...-....)	fre	CH	19XX				Juriste.	Les obligations alimentain JACCOTTET Catherine	1 l.	

Résultat après vérifications des alignements

- Durée : 9h, avec 3 personnes, pour les *multi results*
- Notices d'autorité alignées (*match*) : 1'337
- Notices d'autorité non-alignées (*mismatch, null*) : 1'947
- Total de notices d'autorité traitées : 3'284
- Soit 41 % de notices bénéficient des données d'IdRef

Workflow : de la vérification aux deux bots de saisie



Botorité 2.0

Zone d'identification

Notices d'autorité

Type d'entité	Personne
Forme autorisée du nom	ALEIXANDRE Vicente

Zone de description

Dates d'existence	1898-1984
Historique	Poète espagnol.

Zone du contrôle

Identifiant de notice d'autorité	aleixandre-vicente_idref-026648261
Entretenu par	Centre des littératures en Suisse romande
Statut	Final
Dates de production, de révision et de suppression	Notice ABES. Date de création : 1984-11-09 - Date de modification : 2024-06-26 à 15:00:55.
Sources	Agence bibliographique de l'Enseignement supérieur (ABES) - Données originales récupérées sur http://www.idref.fr/026648261/id .
Notes de maintenance	Saisie automatique [L04] effectuée le : 2024-06-27 à 09:57:15

Crowdsourcer

Zone d'identification

Notices d'autorité

Type d'entité	Personne
Forme autorisée du nom	AGUET Pierre

Zone de description

Dates d'existence	s.d.
Historique	Production participative [crowdsourcing] : Si vous possédez des informations sur cette personne, notamment les dates de naissance et de décès, ainsi que des éléments biographiques, veuillez nous contacter à clsr@unil.ch .

Zone du contrôle

Identifiant de notice d'autorité	aguet-pierre
Entretenu par	Centre des littératures en Suisse romande
Statut	Ébauche
Notes de maintenance	Saisie automatique [L04] effectuée le : 2024-07-01 à 09:28:31

Botorité 2.0 : Saisie automatique



Bilan final

- Bénéfices

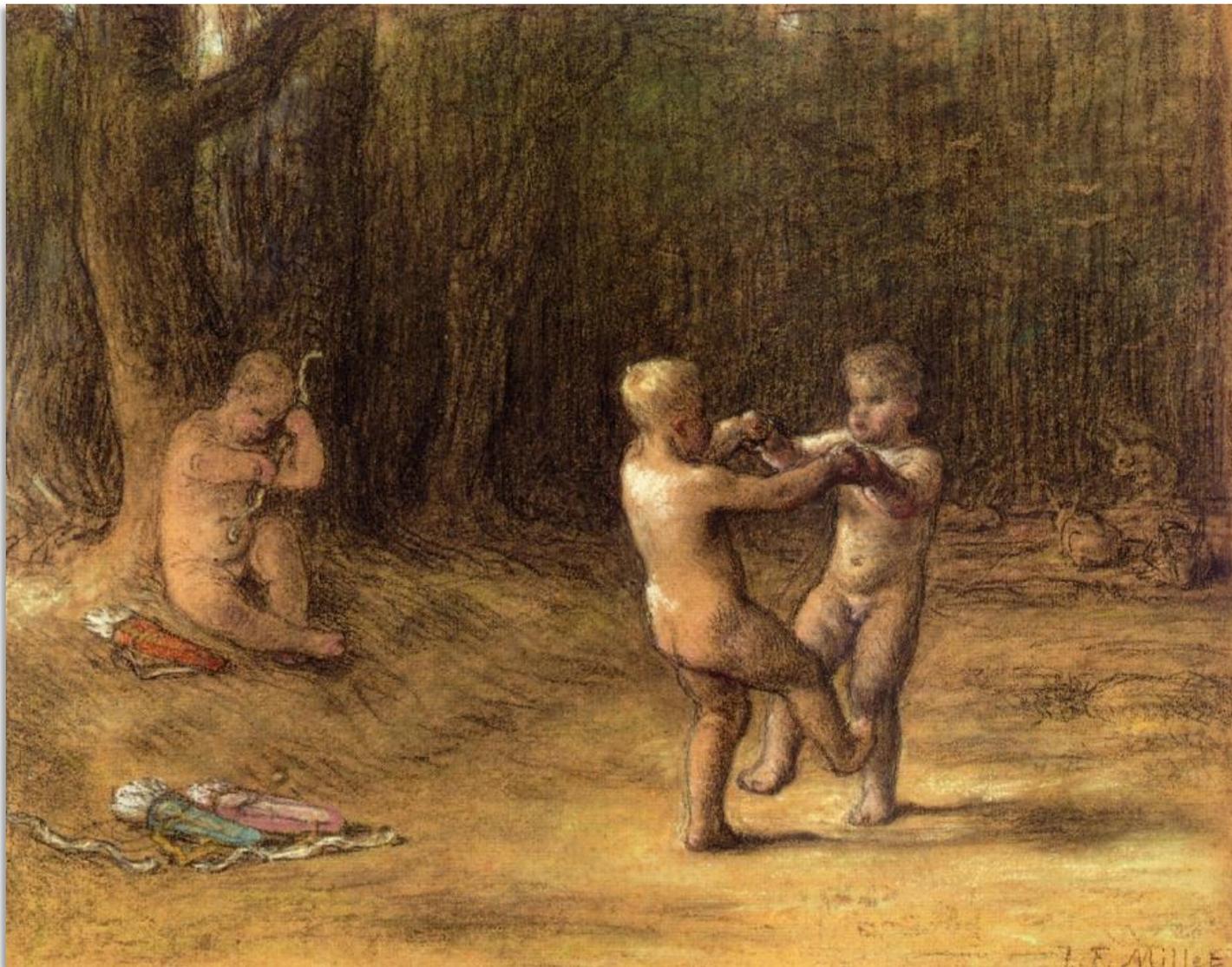
- Amélioration de notre visibilité sur internet
- Enrichissement de notre base de données
- Participation à un réseau professionnel lié aux notices d'autorité
- Mise en lumière de personnes invisibilisées

- Difficultés

- Utilisation de la technologie Python
- Aspect expérimental du projet

- Développements

- Améliorer le flux en utilisant une API ou en réunissant certains scripts, par exemple
- Contribuer à IdRef et à la création de notices d'autorité suisses
- Création d'une *data visualization* cartographiant le réseau de correspondant.e.s



Jean-François Millet, *Danse des Amours*, ca. 1860, fusain et pastel sur papier, Coll. Privée. (source : wikiart.org)

FIN ET REMERCIEMENT

Contact :
nam.pham@unil.ch

CLSR-UNIL



Optimisation du traitement des notices d'autorité dans AtoM et IdRef
© 2024 by Nam Pham is licensed under [CC BY 4.0](https://creativecommons.org/licenses/by-sa/4.0/)