

# Veiller sur les dépôts non affiliés à sa structure dans HAL

Hadrien Finez, Chargé de mission science ouverte, SCD, Université Paris Nanterre

Manon Le Guennec, Responsable des services aux chercheurs, SCD, Université Paris Nanterre  
(<https://orcid.org/0000-0002-0959-9967>)

# Contexte

- Constat : des dépôts mal affiliés (pas toujours affiliés au niveau laboratoire) ou pas affiliés aux structures de recherche de l'UPN qui ne remontaient donc pas dans notre portail UPN
- Dépôts en masse :
  - Par des projets tiers : Migration INRA/IRSTEA, Classiques Garnier, Droit2Hal... Et à l'avenir : Elsevier, Droit2Hal avec d'autres éditeurs ?
- Dépôts par des tiers : auteurs de Nanterre pas toujours affiliés
- Erreurs dans les dépôts : affiliation à des structures non valides
- Volonté de faire de la curation de données et d'enrichir le portail UPN de dépôts qui n'y figuraient pas

# Plusieurs outils

- Un agrégateur de flux RSS : choix de Netvibes
  - Un fichier excel recensant les chercheurs de l'UPN par laboratoire
  - Un fichier word recensant les requêtes API
  - Un tableau de suivi
- Choix de Netvibes :
    - Gratuit
    - Pas de développement à faire
    - Possibilité de travailler sur flux RSS
    - Organisation en tableau pratique en termes de pilotage

# Créer la requête API

[https://api.archives-ouvertes.fr/search/?fq=auth\\_t:\(\"Prénom Nom\"OR\"Prénom Nom\"OR\"Prénom Nom\"\)&fq=structId\\_i:\(\\* NOT \"\[Identifiant HAL de la structure\]\"\)](https://api.archives-ouvertes.fr/search/?fq=auth_t:(\)

- Même avec des guillemets HAL n'est pas sensible à la casse
- L'ordre "Prénom Nom" a son importance
- Si jamais il y a trop de bruit il est possible de filtrer par type "shs" en rajoutant

**&fq=domain\_t:( 0.shs OR 1.shs\* )**

- Ajouter critère &fq=submittedDateY\_i:[X TO \*] (où X représente l'année en cours) afin de filtrer les réponses déjà traitées.

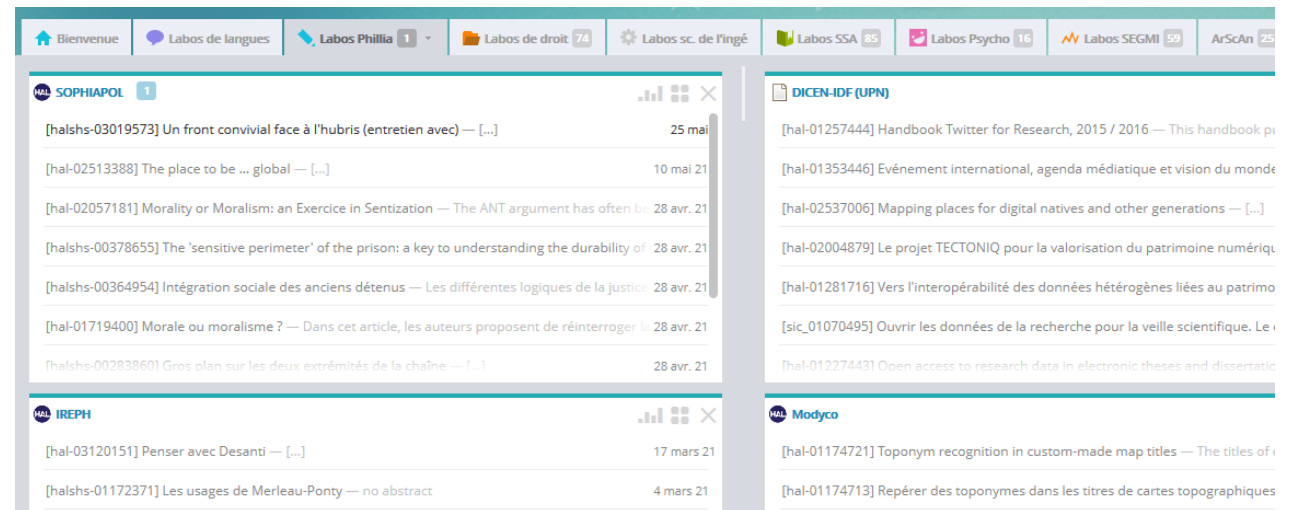
&fq=submittedDate\_s:[AAAA-MM-JJ TO \*] (la date correspond à la dernière vérification)

CREA

[https://api.archives-ouvertes.fr/search/?fq=auth\\_t:\(\"AJI HELENE\"OR\"ARNAUD PIERRE\"OR\"BAISNEE VALERIE\"OR\"BARRY BRIGITTE\"OR\"BASSE JEAN-BERNARD\"OR\"BATESTI ANNE\"OR\"BAUDIER VALERIE\"OR\"BAZIN-FISCHER CLAIRE\"OR\"BIRKS CECILE\"OR\"BOIS CATHERINE\"OR\"BRAILOWSKY YAN\"OR\"BRAUN ALICE\"OR\"BRUNEAU-RUMSEY ANNE-PASCALE\"OR\"BUCKWALTER ELVIS\"OR\"CHATELLIERHUGO\"OR\"CHAUVIN SERGE\"OR\"CHEMALI MARC\"OR\"COULOUMA FLORE\"OR\"CREMIEUX ANNE\"OR\"CROS BERNARD\"OR\"CROWLEY CORNELIUS\"OR\"CRUNELLE ANNY\"OR\"CUSSET FRANCOIS\"OR\"DEBRAS CAMILLE\"OR\"DEVienne ELSA\"OR\"DUBOIS LAURENCE,\"OR\"DUCRAY AMANDINE\"OR\"DUCROCQ MYRIAM-ISABELLE\"OR\"DURIEUX CATHERINE\"OR\"EELLS EMILY\"OR\"ESTRADE CHARLOTTE\"OR\"FOLIARD DANIEL\"OR\"GANAPATHY-DORE GEETHA\"OR\"GERVAIS LAURENCE\"OR\"GHERMANI LAILA\"OR\"HERRY BENIT NADINE\"OR\"HEURTEBIZE FREDERIC\"OR\"JAMET MOREAU EGLANTINE\"OR\"KOHLE R CATHERINE\"OR\"KRAL FRANCOISE\"OR\"LABICA THIERRY\"OR\"LAVABRE ATHENA\"OR\"LE RESTE ANNE-CLAIRE\"OR\"LEROUX-BEAL AGNES\"OR\"LETALLEUR-SOMMER SEVERINE\"OR\"MANOIOV PASCALE\"OR\"MARS LUCA\"OR\"MC CABE JAMES\"OR\"MONFORT BRUNO\"OR\"MUNDLER HELEN\"OR\"MURPHY CHRISTINE\"OR\"PAQUET-DEYRIS ANNE-MARIE,\"OR\"PARAGEAU SANDRINE \"OR\"PAUWELS MARIE CHRISTINE\"OR\"PAYCHA DANIELE\"OR\"PERMUY-MARREC BRIGITTE\"OR\"PHAMKIM-ASSANTE LINH\"OR\"PHILIPPE MURIELLE \"OR\"PRUNIER CLOTILDE\"OR\"RAINERI SOPHIE\"OR\"RAULINE VERONIQUE\"OR\"ROLLAND-DIAMOND CAROLINE\"OR\"ROSSETTE FIONA\"OR\"ROY MICHAEL\"OR\"SANSONETTI LAETITIA\"OR\"SCHNEIDER FLORENCE\"OR\"SEKALI MARTINE \"OR\"SMITH BRADLEY\"OR\"TOTH NAOMI\"OR\"VIOLLAIN CECILE\"OR\"ALEXANDRE GARNER CORINNE\"OR\"BORDAT FRANCIS\"OR\"CHASSAGNOL MONIQUE\"OR\" DECONINCK-BROSSARD FRANCOISE\"OR\"DELOURME CHANTAL\"OR\"DEYSINE ANNE\"OR\"GUERLAIN PIERRE\"OR\"PEDOT RICHARD\"OR\"PERRIN-CHENOUR MARIE-CLAUDE\"OR\"ROY GINETTE\"OR\"SIPERE DOMINIQUE\"\)&fq=structId\\_i:\(\\* NOT \"92837\"\)&rows=1000&wt=rss](https://api.archives-ouvertes.fr/search/?fq=auth_t:(\)

# Insérer la requête API dans le Netvibes

- Options > Ouvrir lien sur site externe
- Attention : pour certains labos avec de nombreux chercheurs, il peut être nécessaire de couper la requête.



The screenshot displays a Netvibes dashboard with a navigation bar at the top containing several lab icons: Bienvenue, Labos de langues, Labos Phillia (1), Labos de droit (74), Labos sc. de l'ingé, Labos SSA (85), Labos Psycho (16), Labos SEGMI (59), and ArScAn (25). The main content area is divided into several panels, each representing a different lab:

- SOPHIAPOL** (1):
  - [halshs-03019573] Un front convivial face à l'hubris (entretien avec) — [...] 25 mai 21
  - [hal-02513388] The place to be ... global — [...] 10 mai 21
  - [hal-02057181] Morality or Moralism: an Exercise in Sentization — The ANT argument has often be 28 avr. 21
  - [halshs-00378655] The 'sensitive perimeter' of the prison: a key to understanding the durability of 28 avr. 21
  - [halshs-00364954] Intégration sociale des anciens détenus — Les différentes logiques de la justice 28 avr. 21
  - [hal-01719400] Morale ou moralisme ? — Dans cet article, les auteurs proposent de réinterroger l. 28 avr. 21
  - [halshs-002838601] Gros plan sur les deux extrémités de la chaîne — [...] 28 avr. 21
- IREPH**:
  - [hal-03120151] Penser avec Desanti — [...] 17 mars 21
  - [halshs-01172371] Les usages de Merleau-Ponty — no abstract 4 mars 21
- DICEN-IDF (UPN)**:
  - [hal-01257444] Handbook Twitter for Research, 2015 / 2016 — This handbook pr
  - [hal-01353446] Événement international, agenda médiatique et vision du monde
  - [hal-02537006] Mapping places for digital natives and other generations — [...]
  - [hal-02004879] Le projet TECTONIQ pour la valorisation du patrimoine numériq
  - [hal-01281716] Vers l'interopérabilité des données hétérogènes liées au patrimo
  - [sic\_01070495] Ouvrir les données de la recherche pour la veille scientifique. Le
  - [hal-012274431] Open access to research data in electronic theses and dissertati
- Modyco**:
  - [hal-01174721] Toponym recognition in custom-made map titles — The titles of
  - [hal-01174713] Repérer des toponymes dans les titres de cartes topographiques

# Traiter une notice qui remonte via la requête API

## **Correction des notices :**

- Correction de l'affiliation
- Enrichissement des métadonnées (au moins DOI si existant)

## **Les cas de « faux positif » :**

- Il s'agit d'un homonyme ( <https://hal.archives-ouvertes.fr/hal-01113963> )
- Il s'agit bien d'une publication du chercheur mais il n'était pas affilié à la structure (<https://halshs.archives-ouvertes.fr/halshs-02888938>)
- Il s'agit bien d'une publication du chercheur mais il a quitté la structure : Mise à jour de la requête API et de la liste des chercheurs

# Limites

- Netvibes :
  - Tableaux peu paramétrables, impossible d'écarter les faux positifs sans modifier la requête API
  - Doit être nettoyé régulièrement pour que ça ne soit pas chronophage
  - Caractères limités (~2000)
- Requête API
  - Il faut repartir à 0 à chaque fois que l'on rajoute des auteurs
  - Ne prend pas en compte les différentes formes auteurs (ex: prénom en initiales)
- Fichier de recensement
  - Doit impérativement être à jour pour être efficace

# Bilan : notices corrigées

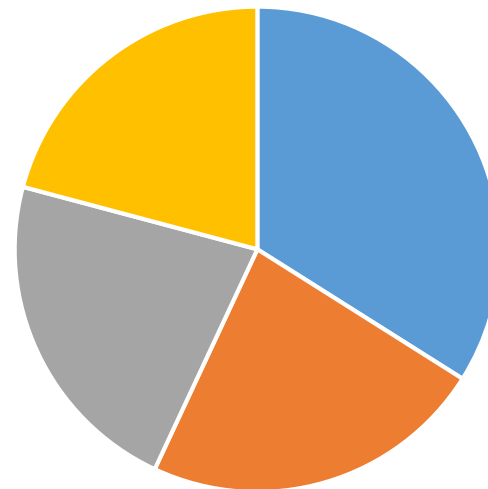
- Année 2019 :
  - 199 notices corrigées dont 22 fichiers
- Année 2020 :
  - 374 notices corrigées dont 84 fichiers
- Année 2021 :
  - 103 notices corrigées
  - 22 fichiers et 43 notices sans TI ont pu être ajoutés au portail



# Bilan 2020

- Année 2020 : principalement des oublis de double affiliation (127), aucune affiliation (86), affiliation incorrecte (83), affiliation institution (78)

Répartition des erreurs d'affiliation

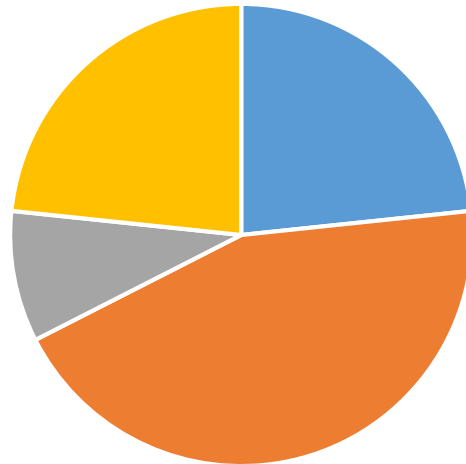


■ Oubli (double affiliation) ■ Absence d'affiliation ■ Affiliation incorrecte ■ Affiliation institution

# Bilan 2021

- Année 2021 : Absence d'affiliation (53), affiliation institution (28), oubli de double affiliation (11), mauvaise affiliation (11)

Répartition des erreurs d'affiliation



■ Oubli (double affiliation) ■ Absence d'affiliation ■ Affiliation incorrecte ■ Affiliation Institution

# A vous de jouer

- Choisissez une structure de recherche de votre établissement
- Créez la requête API pour 10 chercheur(se)s
- Faites-apparaître les résultats sous la forme de votre choix (Netvibes, XML...)
- Corrigez les dépôts pour lesquelles l'affiliation fait défaut

Rappel de la requête API :

```
https://api.archives-ouvertes.fr/search/?fq=auth_t:("Prénom  
Nom"OR"Prénom Nom"OR"Prénom Nom")&fq=structId_i>(*  
NOT "[Identifiant HAL de la structure]")
```