

ExtrHAL – Guide d'utilisation

ExtrHAL est un programme PHP conçu et développé par Olivier Troccaz (ECOBIO/OSUR) et [Laurent Jonchère](#) (Université de Rennes 1), à partir du script ExtractionHAL développé par [Philippe Gambette](#) (UPEM).

Important : ExtrHAL est à l'origine un outil développé spécifiquement pour Rennes 1. Nous avons fait le choix de le mettre gratuitement à disposition de la communauté. Pour autant, **nous ne fournissons pas de support** hors du périmètre des laboratoires affiliés à Rennes 1, mais restons bien sûr à l'écoute de vos suggestions et retours.

Si vous êtes développeur, n'hésitez pas à contribuer au projet ExtrHAL et à partager vos développements.

Table des matières

Installer ExtrHAL.....	1
Prise en main.....	2
1. Point d'entrée.....	3
2. Souligner les auteurs de son labo.....	3
Souligner les auteurs à partir d'une interrogation par idHAL.....	5
3. Choisir la période.....	6
4. Choisir les types de publications.....	6
Cas particuliers (autres publications, etc.).....	7
5. Hcéres : « 20 % les plus significatifs ».....	8
6. Publications inter-équipes ou inter-unités.....	9
7. Autres options.....	10
8. Options et styles de citations.....	10
9. Affichage des résultats.....	11
10. Export des données.....	12
11. ExtrHAL via OCdHAL.....	13

Installer ExtrHAL

Une dizaine d'institutions ont installé ExtrHAL.

Nous recommandons aux utilisateurs non-Rennes 1 d'installer le script ExtrHAL sur le serveur de leur établissement. Ceci vous permettra de personnaliser l'outil et de l'adapter au contexte de votre institution.

Un fichier d'installation est disponible à cette adresse :

<https://halur1.univ-rennes1.fr/ExtrHAL.7z>

Ainsi qu'une documentation :

<https://halur1.univ-rennes1.fr/installation/Installer-ExtrHAL.pdf>

On peut également suivre les dernières modifications sur GitHub :
<https://github.com/OTroccaz/ExtrHAL>

L'outil WinMerge permet de comparer les fichiers et de repérer les lignes à modifier :
<http://winmerge.org/?lang=fr>

Alternativement, si vous avez les compétences, vous pouvez passer directement par Git / GitHub pour gérer le versionnage et faciliter les mises à jour.

Problèmes éventuels liés à l'installation : il peut s'agir de restrictions propres au serveur de votre institution. Exemple : restrictions sur le serveur qui interdisent à l'application ExtrHAL de faire les appels extérieurs vers HAL.

Prise en main

Le principe d'ExtrHAL est très simple : à partir d'un code collection HAL, d'un idHAL ou d'une référence interne HAL, on extrait la liste des publications correspondante, dans le format souhaité.

Exemple : je souhaite extraire la liste des publications de l'unité Inserm U1085 Irset.

1) Je saisis le code collection du laboratoire dans HAL (toujours en majuscules !) :

Code collection HAL (qu'est-ce que c'est ?) :

2) Dans l'onglet « choix des listes », je choisis la période à afficher :

Cliquez sur les titres des menus pour afficher les choix et options

Choix des listes de publications à afficher :

Période : Du Jusqu'au

3) Je sélectionne ensuite les types de publications (obligatoire !) :

(sélection/désélection multiple en maintenant la touche 'Ctrl' (PC) ou 'Pomme' (Mac) enfoncée):

Articles de revue : Tous les articles (sauf vulgarisation)
 Articles de revues à comité de lecture
 Articles de revues sans comité de lecture

4) J'ajoute si nécessaire des options d'affichage :

Options d'affichage et d'export :

5) Je valide

Valider

6) La liste des résultats s'affiche.

Je peux l'extraire en CSV, RTF (Word / LibreOffice), BibTeX :

[Exporter les données affichées en RTF, en CSV, en Bibtex ou VOSviewerDOI](#)

Revoyons à présent chacune des étapes en détail.

1. Point d'entrée

On peut interroger à partir d'un code collection HAL, d'une référence interne (= champ référence interne d'un dépôt HAL) ou d'un idHAL (identifiant unique HAL créé par l'auteur).

Il est possible de combiner code collection et référence interne pour n'afficher que les publications de la collection ayant la valeur renseignée dans ce champ.

En revanche, on ne peut pas combiner collection / référence interne et liste d'idHAL. Mais il est toujours possible de créer une collection ad hoc à partir d'une liste d'idHAL.

Le code collection est toujours en majuscules.

Interrogation idHAL : entrez la chaîne de caractères (ex : olivier-troccaz)

Si vous interrogez plusieurs idHAL : mettez entre parenthèses et remplacez les guillemets par %22 et les espaces par %20.

Exemple : (%22laurent-jonchere%22%20OR%20%22olivier-troccaz%22).

2. Souligner les auteurs de son labo

Code collection HAL pour la liste des auteurs à mettre en évidence (exemple) :

Il est possible de mettre en évidence les auteurs d'une collection (ou l'auteur correspondant à l'idHAL).

Attention : pour un code collection, cette fonctionnalité n'est pas activée à partir des données du référentiel HAL, mais **sur la base d'une liste interne en csv**.

- **Si vous êtes membre d'un labo Rennes 1** : contactez Laurent Jonchère (laurent.jonchere [at] univ-rennes1.fr) si les auteurs de votre unité ou de votre équipe ne sont pas correctement soulignés.
- **Si vous êtes utilisateur non-Rennes 1 et utilisez la version ExtrHAL de Rennes 1** : il est nécessaire de suivre la procédure qui suit : pré-chargez un fichier csv ou txt contenant la liste des auteurs à souligner, en utilisant le modèle proposé dans le menu ExtrHAL :

Code collection HAL pour la liste des auteurs à mettre en évidence (exemple) :

Attention ! Ce champ ne fonctionne que pour les unités affiliées à Rennes 1. Extérieurs à Rennes 1, vous avez la possibilité de mettre en évidence les auteurs de votre collection ou de votre référence interne en pré-téléchargeant un fichier CSV ou TXT réalisé selon ce modèle.

Note : doivent être renseignés les champs suivants : nom, prénom, code collection, code collection équipe (si applicable), arrivée, départ. S'il n'y a pas de valeur à indiquer, la remplacer par un « x » minuscule. Aucun champ ne doit être vide (Entrez x minuscule dans les cellules sans données). Les champ départ / arrivée sont au format AAAA.

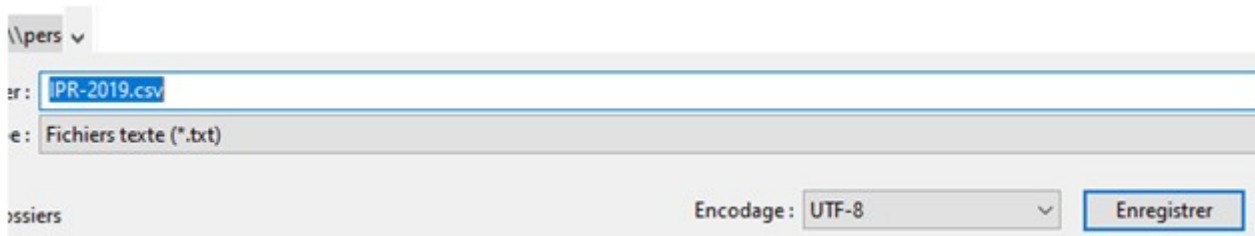
	A	B	C	D	E	F	G	H	I	J
1	Nom	PreNom	Secteur	Titre	Unite	UMR	Grade	Numeq	Eqrec	Collection HAL
2	Fonfec	Sophie	facultatif	facultatif	facultatif	facultatif	facultatif	facultatif	facultatif	Code collection
3	Tim	Vincent	facultatif	facultatif	facultatif	facultatif	facultatif	facultatif	facultatif	Code collection
4	Truand	Tony	facultatif	facultatif	facultatif	facultatif	facultatif	facultatif	facultatif	Code collection
5										
6	NOTE: champs Arrivee / Depart obligatoires, format YYYY ou x minuscule si aucune date									
7	ATTENTION : AUCUN champ ne doit rester VIDE ! Entrer x minuscule dans les cellules sans données									

Une fois votre fichier CSV complété, ouvrez le avec le bloc-notes pour vérifier que vos données sont bien séparées par des **points-virgules** (important) et encodées en **UTF-8** pour éviter les problèmes d'accents (sinon enregistrez-le à nouveau en UTF-8).

IPR-2019.csv - Bloc-notes

Fichier Edition Format Affichage Aide

```
Nom;Prénom;Secteur;Titre;Unité;UMR;Grade;Numeq;Eqrec;Colle
Ababou-Girard;Soraya;SDLM;x;Institut de Physique de Rennes
Alouini;Mehdi;SDLM;x;Institut de Physique de Rennes (IPR);
```



Puis chargez votre fichier dans ExtrHAL (« pré-téléchargez un fichier CSV ») et lancez votre requête.

OU

Si vous avez installé ExtrHAL sur votre serveur : Le fichier [ExtractionHAL-liste-auteurs.php](#) permet de souligner de ou mettre en gras le nom des auteurs de la collection interrogée. Il est possible de modifier et de mettre directement à jour la liste à partir de cette url : <https://url-de-votre-site/ExtractionHAL-liste-auteurs.php>

ExtrHAL : création des fichiers de liste d'auteurs à mettre en évidence

Toutes les collections ▾
Valider

[Ajouter un auteur](#) - [Exporter la liste au format CSV](#) - Importer une liste **entière** ou compléter avec **une liste partielle** à partir d'un fichier csv ou txt ([cf. modèle](#)) -

Total de 2500 auteurs renseignés, toutes collections confondues

Nom	Prénom	Secteur	Titre	Unité	UMR	Grade	Numeq	Eqrec
-----	--------	---------	-------	-------	-----	-------	-------	-------

Important :

- Attention pour l'upload des listes d'auteurs (importer une liste) : après import de la liste complète, faites des sauvegardes en CSV et importez de préférence des **listes partielles** (= une collection, par exemple).
- **Importer une liste entière écrase toute la liste existante** (= toutes les collections) !

Souligner les auteurs à partir d'une interrogation par idHAL

- pour une requête sur un seul idHAL : remplacez les espaces du prénom ou du nom par des tirets bas « _ » (underscore). Exemple : Jean-Luc Le_Breton
- pour une requête sur plusieurs idHAL : séparez en plus les auteurs par un tilde ~ Exemple : laurent-jonchere~olivier-troccaz

Identifiant alphabétique auteur HAL (*idHAL > olivier-troccaz, par exemple*) (Pour une requête sur plusieurs idHAL) : %22laurent-jonchere%22%20OR%20%22olivier-

Auteur correspondant à l'idHAL à mettre en évidence (Instructions) : laurent-jonchere~olivier-troccaz

Dans le menu ExtrHAL (onglet « Options d'affichage et d'export »), la mise en évidence des auteurs du labo interrogé est coché par défaut :

Auteurs (de la collection) ou auteur IdHAL : soulignés gras aucun

Mettre en évidence (\labo{xxxxx}) les auteurs de la collection dans l'export Bibtex : oui non

3. Choisir la période

Saisissez la période au format JJ/MM/AAAA ou recherchez la date dans le calendrier :

The screenshot shows a search form with the following elements:

- Période :** Du Jusqu'au
- Date de dépôt :** Du Jusqu'au
- Articles de revue :** (sélection/désélection)

A calendar dropdown is open, showing the month of June 2015. The calendar has a header with navigation arrows and the text "Juin 2015". The days of the week are abbreviated as L, M, M, J, V, S. The dates 1 through 30 are listed in a grid. Below the calendar, there is a search bar with the text "Recherche 'Ctrl' (PC) ou 'Pomme' (Mac) enfoncée:" and a list of search results including "ation)", "e lecture", "é de lecture", and "nales".

Il est également possible de rechercher ou de filtrer par la date de dépôt dans HAL :

The screenshot shows a search form with the following elements:

- Date de dépôt :** Du Jusqu'au
- Articles de revue :** (sélection/désélection)

A calendar dropdown is open, showing the month of December 2019. The calendar has a header with navigation arrows and the text "Decembre 2019". The days of the week are abbreviated as L, M, M, J, V, S. The dates 1 through 31 are listed in a grid. Below the calendar, there is a search bar with the text "Recherche 'Ctrl' (PC) ou 'Pomme' (Mac) enfoncée:" and a list of search results including "ation)", "e lecture", "é de lecture", "nales", "nales à comité de lecture", and "à comité de lecture".

Note concernant les communications avec actes : ExtrHAL affichera la date du colloque comme date de publication, et non la date de publication des actes. La raison en est que le champ utilisé pour la date de publication dans les communications n'est pas le même que pour les articles de revues (producedDate). Autrement dit, même si la date de publication est renseignée dans HAL, il n'y a pas de champ API HAL permettant de la récupérer. Cette lacune a été signalée au CCSD en juillet 2019 qui a répondu [ticket #74242]: «Ce problème nous a déjà été remonté. Il sera traité prochainement par notre équipe technique. »

4. Choisir les types de publications

Sélectionnez les types de publications que vous souhaitez exporter.

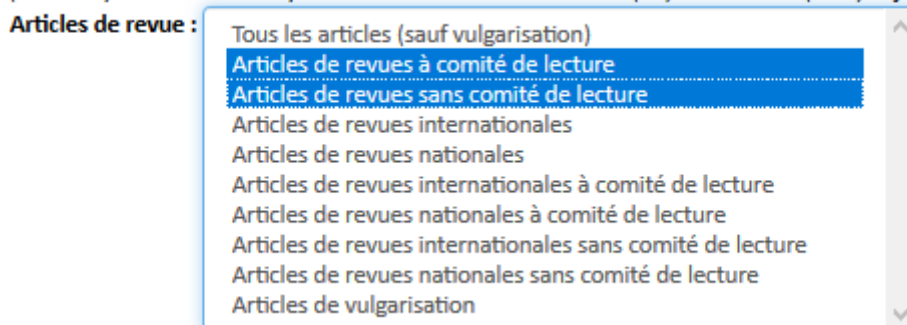
ExtrHAL les regroupe en 4 grandes catégories :

- Articles de revue
- Communications (congrès, conférences, posters)

- Ouvrages, directions d'ouvrages, chapitres, coordinations de numéro de revue
- Autres publications : rapports, brevets, preprints, etc.

Il est possible de sélectionner plusieurs types au sein de ces 4 ensembles :

(sélection/désélection multiple en maintenant la touche 'Ctrl' (PC) ou 'Pomme' (Mac) enfoncée):



Attention : la qualité de l'extraction dépend bien sûr de la qualité des dépôts HAL de la collection. Si l'on souhaite par exemple extraire les articles de revues sans comité de lecture, il faut bien s'assurer que les dépôts HAL correspondants contiennent la bonne information (champ « comité de lecture » = « Non »). De même pour l'audience (Nationale / Internationale) et les communications (avec ou sans actes, invitées ou non invitées).

Pour ces vérifications, [l'application OcdHAL](#) développée par Patricia Reynier et Pascal Bellemain s'avère des plus utiles.

Vulgarisation	Non
Comité de lecture	Oui
Invité	Non
Audience	Internationale
Actes	Non

Cas particuliers (autres publications, etc.)

Pour une documentation plus complète, voir le document produit par l'Université de Toulouse (F. Gouzi) [diffusé sur notre site web open access](#).

Cela concerne les types de documents suivants :

- Coordination de numéro de revue
- Article de blog
- Compte rendu d'ouvrage ou note de lecture
- Notice d'encyclopédie ou de dictionnaire
- Traduction

5. Hcéres : « 20 % les plus significatifs »

Hcéres : « numéroté les publications de 1 à x dans l'ordre chronologique, en distinguant les 20% les plus significatifs puis les 80% restant. »

<https://www.hceres.fr/fr/guides-des-produits-de-la-recherche-et-activites-de-recherche>

1. **Faites une 1ère requête** sur la collection de votre choix, en cochant dans le menu "Identifiant HAL = visible",
2. exportez les résultats en CSV, puis faites dans le fichier CSV une sélection des 20%,
3. créez un nouveau CSV avec une seule colonne contenant les identifiants HAL sélectionnés, avec en-tête "HalId" (sur le modèle proposé dans le menu).
4. **Faites une 2e requête**, en cochant bien "Identifiant HAL = visible" (par défaut) et "Mise en évidence des notices les plus significatives = visible",
5. chargez votre fichier CSV d'identifiants HAL en cliquant sur le lien "lancer la procédure d'extraction à partir de votre liste CSV",
6. validez puis exportez.

	A
1	HalId
2	cea-01988433
3	hal-01619411
4	hal-01619412
5	hal-01638294
6	hal-01639691
7	hal-01655640
8	hal-01660838
9	hal-01671626
10	hal-01671628
11	hal-01685672
12	hal-01685689
13	hal-01685695
14	hal-01695548
15	hal-01695550

Les publications sont matérialisées par une petite flèche :

2019

→ 1. [Cadio J. S., Al Tabaa O., Nguyen C.-D., Faccin M., Guillin R. et al. \(2019\). Back pain following instillation](#)
- /10.1007/s10067-019-04500-w. Réf. HAL: [hal-02087901](#)

2. [Cattoir V., Felden B. \(2019\). Future Antibacterial Strategies: From Basic Concepts to Clinical Challenges.](#)

3. [Kang K. N., Klein D. R., Kazi M. I., Guerin F., Cattoir V. et al. \(2019\). Colistin heteroresistance in Enteroba](#)

Attention : le lien raccourci bitly ne permet pas de sauvegarder le tri des 20%

Plusieurs formats de sortie sont proposés dans le menu :

Mise en évidence des notices les plus significatives (*il est nécessaire de lancer la procédure d'extraction à partir de votre liste CSV réalisée selon ce modèle*) :

visible (→)

visible (20% / 80 %)

visible (20%)

6. Publications inter-équipes ou inter-unités

ExtrHAL permet de repérer, dans la liste des résultats, les différentes collections associées à une publication, et le cas échéant d'en déduire des collaborations entre équipes ou unités de recherche.

Pour activer cette fonctionnalité, il est nécessaire de renseigner dans le menu (« options d'affichage et d'export ») les codes collections des équipes ou unités à remonter :

Numérotation/codification par équipe :

oui

. Nombre d'équipes :

. Nom HAL équipe 1 :

. Nom HAL équipe 2 :

. Nom HAL équipe 3 :

. Nom HAL équipe 4 :

. Limiter l'affichage seulement aux publications croisées : non oui

Liste des résultats :

2018

1. IPR-MN - 1 - IPR-MD - 1 - Davoy X., Gellé A., Lebreton J.-C., Tabuteau H., Soldera A. *et al.* (2018). High Water Flux w
2. IPR-MN - 2 - IPR-MM - 1 - Goujon F., Bêche B., Malfreyt P., Ghoufi A. (2018). Radial-based tail methods for Monte C
3. IPR-MN - 3 - IPR-MM - 2 - Li Q., Garnier L., Vié V., Lhermite H., Moréac A. *et al.* (2018). Sphingolipid Gel/Fluid Phasi

La publication 3 est la 3^e de la collection IPR-MN, et la 2^e de la collection IPR-MM.

7. Autres options

Il est théoriquement possible d'afficher l'IF (impact factor) de la revue pour chaque publication. Nous n'encourageons pas spécialement le recours à cette fonctionnalité car :

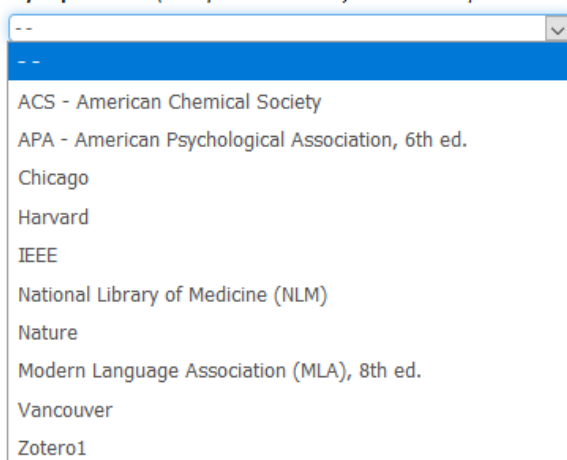
- Faute d'ISSN dans la liste des JIF, ExtrHAL fait une recherche sur le titre des revues. Si la correspondance entre le titre de la liste des JIF et le titre HAL n'est pas exacte, l'IF ne sera pas affiché.
- L'IF est calculé pour une période donnée. Or, le fichier JIF utilisé par ExtrHAL concerne une année spécifique. Il ne peut donc pas être appliqué à n'importe quelle période interrogée.
- L'IF est-il un indicateur pertinent pour l'évaluation ? Il serait plus intéressant d'afficher le nombre de téléchargements du fichier HAL, mais ce champ n'est actuellement pas interrogeable via les webservice HAL.

8. Options et styles de citations

Attention : cette fonctionnalité est expérimentale et concerne essentiellement les articles de revues.

Styles prédéfinis : on peut choisir une norme de citation parmi une liste restreinte.

Styles prédéfinis : (l'adéquation avec le style demandé dépend des éléments qui ont été renseignés dans HAL)



Exemples :

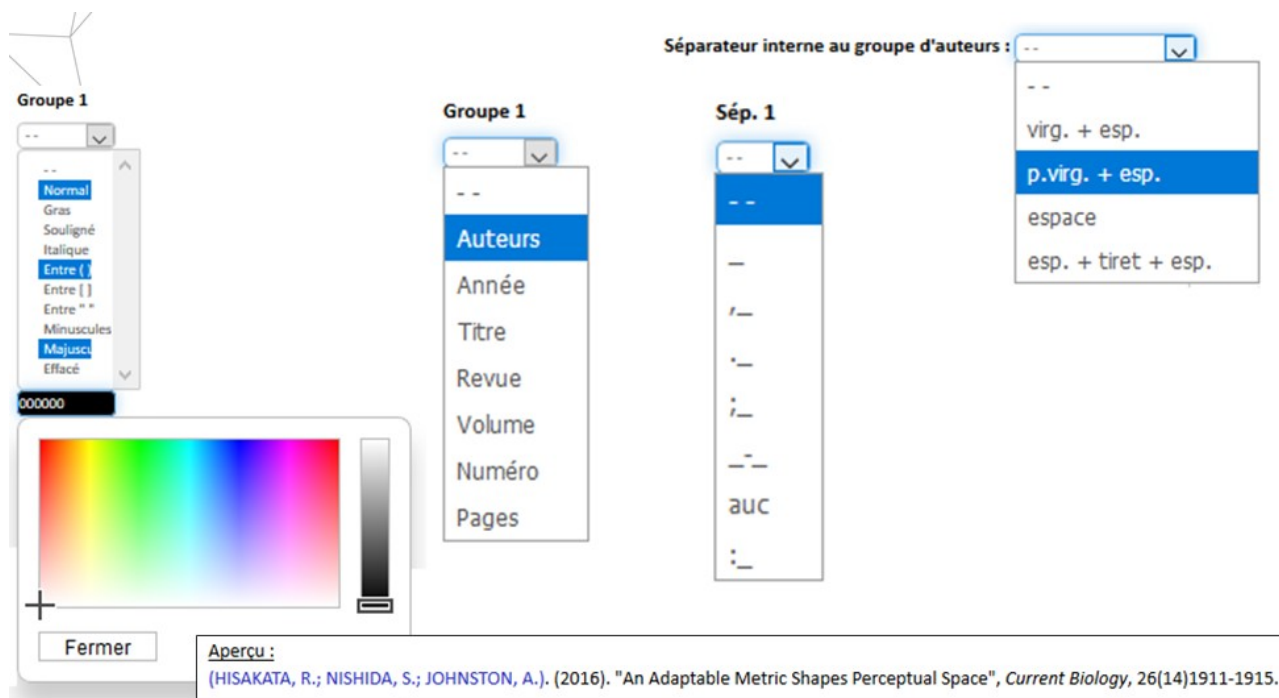
Aperçu :

x. Hisakata, R., Nishida, S., Johnston, A. (2016). An Adaptable Metric Shapes Perceptual Space. *Current Biology*, 26(14), 1911-1915

(x) Hisakata, R.; Nishida, S.; Johnston, A. An Adaptable Metric Shapes Perceptual Space. *Current Biology*. 2016, 26 (14), 1911-1915

Hisakata, Rumi, et al. "An Adaptable Metric Shapes Perceptual Space". *Current Biology*, vol. 26, no. 14, 2016, pp. 1911-1915.

Styles personnalisés : vous définissez vous-même la métadonnée à afficher (auteur, titre, etc.) dans la liste déroulante, ainsi que l'ordre des métadonnées et les séparateurs, voire même les couleurs de police.



9. Affichage des résultats

Une fois la requête validée, la liste est affichée au format HTML. Vous pouvez naviguer dans la liste des résultats à partir du sommaire.

L'URL de la requête API est proposée pour chaque type de document (format JSON) :

Articles de revues à comité de lecture

[URL requête API HAL](#)

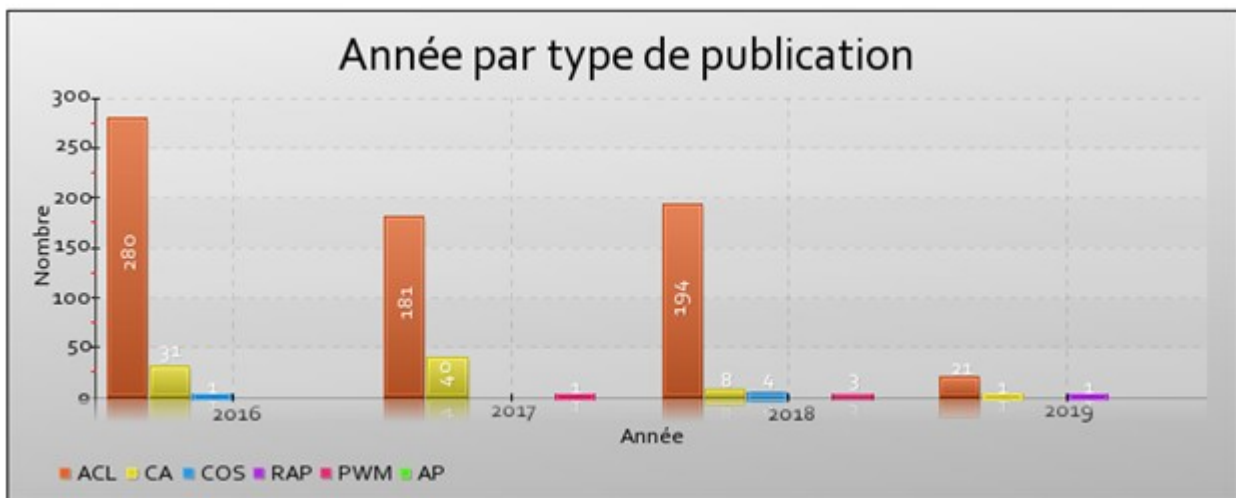
Il est possible d'afficher un lien raccourci (bitly) vers la requête, en cochant l'option « Disposer d'une URL raccourcie directe » du menu ExtrHAL.

Un bilan quantitatif est proposé, accompagné d'une légende :

Bilan quantitatif

	2016	2017	2018	2019
ACL	280	181	194	21
CA	31	40	8	1
COS	1	0	4	0
RAP	0	0	0	1
PWM0	1	3	0	0
AP	0	0	0	0
Total	312	222	209	23

Ainsi que des graphes :



10. Export des données

[Exporter les données affichées en RTF](#), [en CSV](#), [en Bibtex](#) ou [VOSviewerDOI](#)

4 formats d'exports :

- **RTF** (Word, LibreOffice, etc.),
- **CSV** (tableur : Excel, Calc...) : on peut spécifier le séparateur dans le menu : « Délimiteur export CSV » (point-virgule, £ ou paragraphe)
- **BibteX** : ce format a fait l'objet d'une révision approfondie récente dans ExtrHAL (NB : s'agissant des directions d'ouvrages, la nomenclature HAL ne permet pas de distinguer auteurs / éditeurs en sortie BibteX, il faut donc corriger les notices concernées après export.),
- **VOSviewer** : exporte une liste de DOI au format texte pour analyse dans l'outil de visualisation [VOSviewer](#). Alternativement, il est possible d'importer le fichier JSON de la requête ExtrHAL en faisant un clic droit sur « URL requête API HAL » :

Tous les articles (sauf vulgarisation) ↑

[URL requête API HAL](#)

2019

1. Aissani L., Lacassi

2. Atlan A., Van Tilb

3. Bahers J.-B., Giac

4. Bailleul H., Baud

- Ouvrir le lien dans un nouvel onglet
- Ouvrir le lien dans un nouvel onglet contextuel >
- Ouvrir le lien dans une nouvelle fenêtre
- Ouvrir le lien dans une fenêtre de navigation privée
- Marque-page sur ce lien
- Enregistrer la cible du lien sous...

11. ExtrHAL via OCdHAL

Vous pouvez également utiliser la version ExtrHAL intégrée à l'outil [OcdHAL](#) (Patricia Reynier / Pascal Bellemain).

Spécificités de la version OCdHAL :

- Pas besoin d'installation,
- Export de toute la collection ou d'une sélection,
- Export via un identifiant structure ou un code collection,
- Mise en évidence des auteurs de la collection (sans passer par une liste CSV),
- Sélection des 20% via l'interface (champ Hcéres),
- L'utilisateur peut définir sa propre nomenclature,
- Aperçu immédiat des résultats.

Une documentation complète et détaillée est proposée dans le [Manuel d'utilisation OcdHAL](#) (p. 39 sq).

Laurent Jonchère / 02-12-2019