



Mesurer les collaborations entre unités de recherche

Un exemple d'utilisation des API de HAL

Didier Vercueil (Observatoire de la recherche - DGD-RIV)
Marion Etasse (mission Outils, Support, Innovations – SID)
Samuel Libot (mission Outils, Support, Innovations – SID)

SOMMAIRE

- Objectif : pourquoi mesurer ces collaborations ?
- Réalisation : comment récupérer les données ?
- Exploitation des résultats : forme et contextualisation

Pourquoi mesurer ces collaborations ?

Objectif et origine de la démarche

- Développer l'interdisciplinarité est un objectif majeur du site (index, UGA)
- L'évolution des mesures des collaborations entre unités (entre pôles) permet une première approche pour évaluer l'interdisciplinarité
- Première approche réalisée en 2016 avec le WOS (manque d'éléments pour les SHS)
- Période favorable : fin dossier évaluation HCERES des unités
- Travaux utilisés par les pôles de recherche pour le travail d'autoévaluation

Récupérer les données

Étape 1 : définir le périmètre de laboratoires à interroger

"périmètre **COMUE**" dans HAL (périmètre de travail de la DGDRIV)

- **labos UGA**
- **labos INP**
- **labos co-tutelle UGA/CEA**
- **labos co-tutelle UGA/CNRS**
- **labos USMB**
- **labos IEP**
- **labos co-tutelle UGA/INRIA**

Récupérer les données

Étape 2 : comment traduire le périmètre défini au préalable dans HAL ?

Dans l'idéal, on voudrait récupérer les ID des labos en se basant sur les ID des tutelles.

Mais certaines structures ne sont pas forcément affiliées correctement. Au fil des tests, on a enrichi la liste des labos avec des **ID directs** en utilisant d'autres sources (RNSR, DGDRIV...)

Requête actuelle pour récupérer les ID des labos dont on veut mesurer la collaboration :

274 ID de labos à croiser.

Récupérer les données

- [https://api.archives-ouvertes.fr/ref/structure/?q=\(\(parentDocid_i:\(445543 OR 51016 OR 3886 OR 5485 \) OR parentDocid_i:\(89889 OR 300275 OR 300339\) OR parentDocid_i:\(445543 AND 441569\) OR \(parentDocid_i:\(301286 OR 300016\) AND parentDocid_i:445543\) OR parentDocid_i:\(445543 AND 2497\) OR parentDocid_i:89636\) OR docid:\(37786 OR 97447 OR 20150 OR 48098 OR 432899 OR 222134 OR 79 OR 129 OR 599 OR 60436 OR 723 OR 60436 OR 157379 OR 7118 OR 14088 OR 85530 OR 93721 OR 596 OR 25119 OR 40221 OR 24549 OR 80446 OR 479681 OR 490594 OR 2497 OR 86972 OR 441128 OR 221029 OR 216283 OR 221028 OR 74115 OR 182213 OR 460264 OR 460265 OR 25118\)\) AND valid_s:VALID\) OR docid:\(732 OR 664 OR 182211 OR 182208\)&wt=json&rows=300&fl=docid,label_s,acronym_s](https://api.archives-ouvertes.fr/ref/structure/?q=((parentDocid_i:(445543 OR 51016 OR 3886 OR 5485) OR parentDocid_i:(89889 OR 300275 OR 300339) OR parentDocid_i:(445543 AND 441569) OR (parentDocid_i:(301286 OR 300016) AND parentDocid_i:445543) OR parentDocid_i:(445543 AND 2497) OR parentDocid_i:89636) OR docid:(37786 OR 97447 OR 20150 OR 48098 OR 432899 OR 222134 OR 79 OR 129 OR 599 OR 60436 OR 723 OR 60436 OR 157379 OR 7118 OR 14088 OR 85530 OR 93721 OR 596 OR 25119 OR 40221 OR 24549 OR 80446 OR 479681 OR 490594 OR 2497 OR 86972 OR 441128 OR 221029 OR 216283 OR 221028 OR 74115 OR 182213 OR 460264 OR 460265 OR 25118)) AND valid_s:VALID) OR docid:(732 OR 664 OR 182211 OR 182208)&wt=json&rows=300&fl=docid,label_s,acronym_s)

Récupérer les données

Étape 3 : comment exprimer et mesurer la collaboration entre deux laboratoires ?

On s'appuie sur les facettes pour voir le nombre de publications en commun entre le labo testé et les autres

https://api.archives-ouvertes.fr/search/saga/?q=structId_i:%22224470%22&wt=json&rows=0&facet=true&facet.field=labStructId_i

Récupérer les données

- Utilisation des API dans script R
- On récupère pour chaque publi les différents id des unités
- On peut créer les matrices de collaborations unités*unités
- On peut analyser ces matrices : évolutions au fil du temps ; intra-pôles ; inter-pôles...

Récupérer les données

- Permet seulement d'évaluer une tendance, car données HAL :
 - Doublons !!!
 - Labos qui renseignent encore peu (mais pour que la collab. ne soit pas prise en compte, il faudrait que les deux labos ne renseignent pas)
 - Travail sur référentiel encore à affiner (notamment CBS avec CHU)
- Problématique des gros labos pluridisciplinaires (GIPSA par exemple) : il faudrait une approche par équipe

Analyse en pôles

Evolution 2015-2016

	PAGE	CBS	MSTIC	SHS	PSS	CEA
PEM	24	-10	12	7	13	19
PAGE		0	5	12	12	4
CBS			0	6	1	2
MSTIC				3	-7	7
SHS					-9	0
PSS						0

Evolution 2016-2017

	PAGE	CBS	MSTIC	SHS	PSS	CEA
PEM	-4	10	0	-6	-5	17
PAGE		-7	-14	1	-3	1
CBS			5	4	1	-6
MSTIC				4	9	-1
SHS					-5	-16
PSS						-1

Evolution 2017-2018

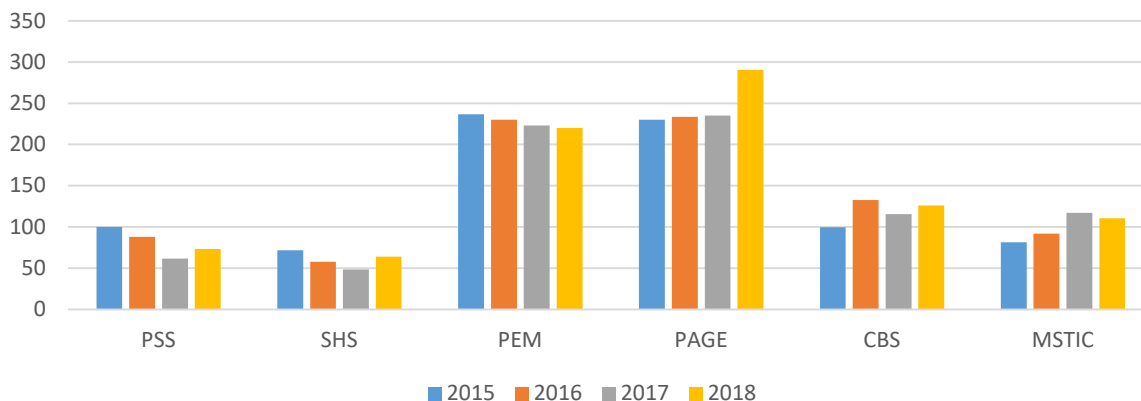
	PAGE	CBS	MSTIC	SHS	PSS	CEA
PEM	-1	-13	7	9	1	-27
PAGE		10	13	-9	-2	4
CBS			4	6	-1	-4
MSTIC				8	-4	1
SHS					4	9
PSS						-1

Exploiter les résultats

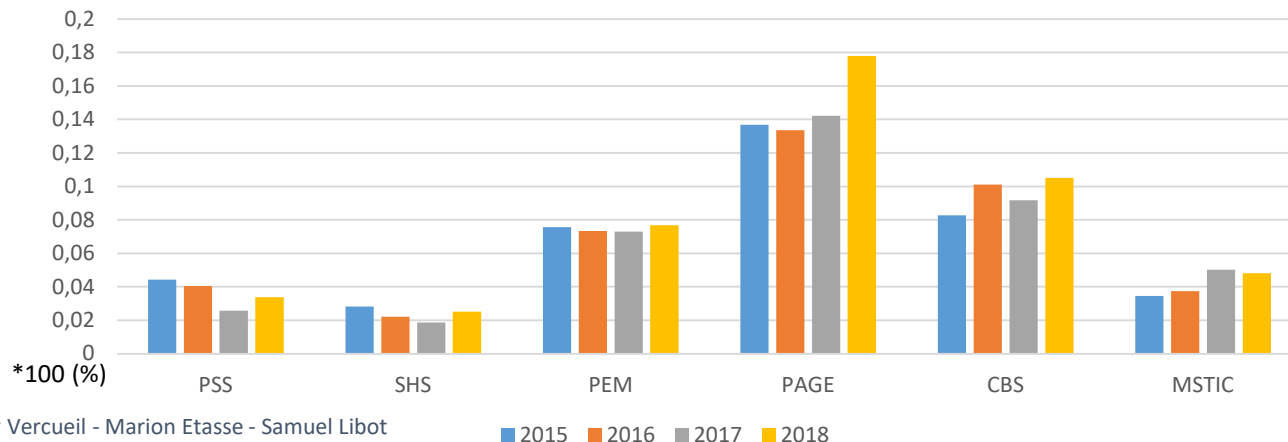
Tendance (analyse en %)

- Augmentation pour PAGE-CBS-MSTIC
- Egale pour PEM
- Baisse puis augmentation en 2018 pour SHS et PSS

Nb de collaborations intra-pôle



Collaboration intra-pôle (en %)



Exploiter les résultats

Visualisation de
graphe entre unités

