

# Caractérisation de la production scientifique des acteurs de la recherche culturelle soutenue par le ministère de la Culture dans le cadre de sa politique de recherche

**RAPPORT AU MINISTERE DE LA CULTURE**

Mai 2018

*Ce rapport a été réalisé pour le Ministère de la culture. L'équipe projet de l'OST était composée de Aouatif de la Laurencie (chef de projet), Anne Glanard (chargée d'études), Abdelghani Maddi (chargé d'études statistiques), Mathieu Goudard (chargé d'études statistiques). Frédérique Sachwald a assuré un suivi de la réalisation du projet et de la rédaction du rapport en tant que directrice de l'OST. Lamy Lakehoul et Nelson Teixeira du département du système d'information du Hcéres ont contribué au calcul des indicateurs issus de la base de publications de l'OST (correspondant au corpus MC-WoS dans le rapport).*

## PRÉAMBULE

Le département de la recherche, de l'enseignement supérieur et de la technologie (DREST) du Ministère de la culture et de la communication (MC) est en charge de la stratégie, du suivi et de la mesure des impacts des activités d'enseignement supérieur et de recherche dans le domaine culturel. Le DREST doit également fournir des informations et des chiffres clés aux différents acteurs (cabinet du ministre, départements du MC, parlement, Conseil national de l'enseignement supérieur et de la recherche artistiques & culturels, ministères,...). Le DREST souhaite accompagner les unités et les établissements du domaine lors de leur évaluation par le Hcéres.

Pour remplir ses missions, le DREST envisage de produire un rapport d'activité avec des chiffres clés sur l'activité d'enseignement supérieur et de recherche dans les établissements sous tutelle ou co-tutelle du MC. Dans cette perspective, il a fait appel à l'OST pour l'accompagner.

Il est convenu que ce dernier accompagnera également le DREST dans la mise en place d'un dispositif de mesure et de suivi de la production scientifique qui couvrira l'enseignement supérieur de culture (ESC) et la recherche culturelle.

L'accompagnement par l'OST est organisé en trois phases.

La phase 1 consiste à expertiser les sources mobilisables sur la production scientifique du MC avec :

- une exploration des bases de données HAL<sup>1</sup> et Web of Science (WoS)<sup>2</sup> ;
- une analyse des listes de publications fournies dans les dossiers d'évaluation reçus par le Hcéres.

Lors de la phase 2, l'OST produit un rapport sur :

- la qualification des différentes sources de données, notamment leur représentativité et la qualité des données dans la perspective d'un suivi de la production scientifique et du calcul d'indicateurs pertinents ;
- les types de production dans les domaines de la recherche culturelle par les entités sous tutelle du MC.

La phase 3 porte sur la proposition d'un dispositif de suivi au DREST afin de poursuivre le travail d'identification et de caractérisation de la production dans les domaines de la recherche culturelle (Partie II).

---

<sup>1</sup> Créé en 2001 par le CNRS, Hyper articles en ligne (ou HAL) est une archive ouverte permettant aux chercheurs français de déposer leurs productions dans une base à accès ouvert développée par le Centre pour la communication scientifique directe (CCSD) du CNRS.

<sup>2</sup> Web of Science (WoS) de Clarivate Analytics, base de données citationnelle internationale

# SOMMAIRE

<b>PARTIE I</b> .....	<b>3</b>
<b>SYNTHÈSE</b> .....	<b>4</b>
<b>INTRODUCTION</b> .....	<b>6</b>
<b>DONNEES ET METHODE</b> .....	<b>7</b>
1 / Sources des données sur la production des structures MC .....	7
<i>Corpus issus du Web of Science et de HAL</i> .....	7
<i>Les données issues des dossiers d'autoévaluation du Hcéres</i> .....	8
2 / Définitions et calcul d'indicateurs .....	9
<b>ANALYSE DU CORPUS MC-WOS</b> .....	<b>10</b>
1 / Production du MC recensée dans le WoS .....	10
2 / Langues des publications .....	11
3 / Domaines de recherche et thématiques du corpus MC-WoS .....	12
4 / Co-publications et réseaux au sein du corpus MC-WoS .....	16
5 / Impact des publications du corpus MC-WoS .....	19
<b>ANALYSE DU CORPUS MC-HAL</b> .....	<b>20</b>
1 / Types de dépôt : archives et notices .....	20
2 / Nombre de dépôts et types de productions du corpus MC-HAL .....	21
3 / Langue de production .....	23
4 / Domaines de recherche et thématiques .....	24
<i>Distribution du corpus par domaine de recherche</i> .....	24
<i>Cartographie des thèmes de recherche dans le corpus MC-HAL</i> .....	27
<b>ANALYSE DU CORPUS MC-HCÉRES</b> .....	<b>28</b>
1 / Nombre et types de production .....	28
2 / Production par type de structure .....	30
3 / Structures et productions par domaine disciplinaire .....	30
4 / Types de production selon les domaines disciplinaires .....	32
<b>COMPARAISON DES TROIS CORPUS</b> .....	<b>33</b>
1 / Types de documents et revues selon les corpus .....	33
2 / Langues de production .....	34
3 / Profil disciplinaire des corpus MC-WoS et MC-HAL .....	34
4 / Indicateurs : spécialisation WoS et HAL ; impact WoS uniquement .....	35
5 / Co-publications et partenariats .....	35
<b>CONCLUSIONS ET APPROFONDISSEMENTS</b> .....	<b>36</b>
<b>REFERENCES</b> .....	<b>38</b>
<b>Annexe 1 - Liste des structures prises en compte</b> .....	<b>39</b>
<b>Annexe II - Table de correspondance « Types de doc /Production » entre sources</b> .....	<b>44</b>
<b>Annexe III - Nomenclature disciplinaire</b> .....	<b>47</b>
<b>Annexe IV – Indice d'impact</b> .....	<b>49</b>
<b>PARTIE II - Dispositif de suivi présenté le 6 avril 2018 au MC</b> .....	<b>50</b>



## **PARTIE I**

## SYNTHÈSE

L'analyse de la production des structures du périmètre du Ministère de la culture s'appuie sur trois sources de données complémentaires : le Web of Science, base bibliométrique internationale permettant de mesurer les citations et les collaborations internationales, l'archive nationale HAL et les listes des productions fournies lors des évaluations par le Hcéres. Trois corpus spécifiques ont été constitués à partir de ces sources : les corpus MC-WoS et MC-HAL portent sur la période 2007-16, soit une dizaine d'années, alors que le corpus MC-Hcéres suivant les vagues d'évaluation, il ne porte sur 5 années glissantes comprises dans la période. Ils présentent chacun des avantages et des inconvénients pour recenser et caractériser la production des unités de recherche. Chacun permet de calculer certains indicateurs et les comparaisons ne sont possibles que pour certaines caractéristiques de la production des structures MC.

### **Volumes de production et types de documents dans les trois corpus**

Le corpus MC-WoS (3 594 publications) est essentiellement composé d'articles parus dans des revues à comité de lecture ou publiés dans des actes de conférences (93%).

Les corpus MC-HAL (29 224 productions) et MC-Hcéres (43 200 productions) sont chacun composés de 4 types de productions : le premier est formé d'articles de revues et de contributions à des congrès avec actes publiés (36% dans les deux cas), puis viennent les ouvrages, les communications à des congrès non publiées et les « autres produits ». Cette dernière catégorie comporte notamment des produits spécifiques de la recherche culturelle (rapports de fouille, catalogue d'exposition...).

En passant de la totalité des corpus au seul périmètre des publications (articles, ouvrages et actes) le ratio entre le nombre de productions recensées dans le WoS et dans HAL passe de 8 à 5, et entre les productions recensées dans le WoS et le corpus Hcéres, de 12 à 6.

### **Langues des documents selon les corpus et les disciplines**

Dans le corpus MC-HAL, l'utilisation du français domine mais l'usage de l'anglais progresse : le profil est le même pour le corpus HAL limité aux domaines SHS (HAL dom. SHS). Dans le corpus MC-WoS à l'inverse, l'anglais domine, ce qui correspond aux pratiques constatées dans les corpus élargis à toutes les disciplines scientifiques.

### **Profil disciplinaire et thématique des productions des structures MC**

Si les classifications varient entre le WoS et HAL, l'analyse permet toutefois d'indiquer que :

- l'archéologie, l'architecture, l'histoire de l'art et l'histoire, ainsi que la géographie et la paléontologie sont des thématiques communes aux 2 corpus ;
- l'anthropologie et les géosciences ont une présence plus marquée dans le corpus MC-WoS ;
- l'informatique et la musique sont plus visibles au sein du corpus MC-HAL.

Les cartographies thématiques, réalisées pour les 2 corpus MC-WoS et MC-HAL, présentent des paysages complémentaires :

- sur MC-WoS, 4 grands groupes thématiques ont été délimités : archéologie/paléontologie, géographie physique, histoire de l'art, et technologies mises en œuvre dans les domaines du MC ;
- sur MC-HAL, en plus de l'archéologie et de la géographie, des thématiques liées à l'urbanisme, l'architecture, la littérature et l'histoire de l'art, ainsi que la musique et la linguistique, apparaissent.

Le positionnement du MC au sein des productions de la recherche française peut être apprécié par l'indice de spécialisation :

- le corpus MC-WoS est fortement spécialisé en archéologie, architecture, anthropologie et histoire de l'art. Il a une activité relativement plus faible en géosciences et géographie physique ; c'est également le cas en paléontologie et histoire ;
- le corpus MC-HAL au sein de l'ensemble des dépôts dans l'archive HAL est spécialisé en informatique/son, archéologie et préhistoire, architecture/ aménagement de l'espace. L'analyse fait ainsi apparaître le domaine de la musique lié à l'informatique alors que la cartographie sur le corpus MC-WoS identifiait plutôt une thématique acoustique.

### **Impact des publications du corpus MC-WoS**

Sur l'ensemble de la période, les publications du corpus MC-WoS reçoivent plus de citations que la moyenne des publications françaises.

L'anthropologie est le domaine de recherche dans lequel le corpus MC a les impacts les plus élevés : ses publications reçoivent 3,3 fois plus de citations par publication que la moyenne des publications mondiale et 1,6 fois plus que les publications françaises dans ce domaine. En archéologie, le nombre de citations par publication est 2,4 fois plus élevé pour le corpus MC-WoS que pour le monde et 1,4 fois plus élevé que pour la France. Dans le domaine de l'architecture, les publications du MC reçoivent plus de citations que les publications françaises(1,4) et un nombre équivalent à celui de l'ensemble des publications mondiales.

En paléontologie, le corpus MC-WoS a un impact proche de la moyenne mondiale, mais inférieur à celui de l'ensemble des publications françaises.

### **Partenaires des structures MC**

Les structures en tutelle ou co-tutelle du MC suivent la tendance générale à l'accroissement de la part des co-publications, toutes disciplines confondues. Le corpus MC- WoS permet de distinguer leurs principaux partenaires à l'international qui sont les États-Unis et les grands pays européens (Espagne, Italie, Allemagne...). Les réseaux de collaboration avec l'ensemble des pays donnent une vision plus complète des partenariats : les structures MC ont, comme dans les autres domaines de recherche, des collaborations plus fréquentes avec les pays d'Europe. Trois grands réseaux de co-publication se dessinent à l'échelle mondiale : l'un centré sur la France, l'autre comprenant l'Italie et le Canada, enfin, celui composé des États-Unis, du Royaume-Uni et de l'Allemagne.

## INTRODUCTION

Ce rapport étudie les productions des structures de recherche soutenues par le Ministère de la culture (MC) afin de les caractériser. Les structures de recherche soutenues par le MC sont soit des unités en co-tutelle avec le CNRS, soit des services à compétence nationale (SCN), soit des établissements publics nationaux (EPN) qui ont également une mission de recherche.

Le périmètre institutionnel MC de l'étude est composé des établissements publics nationaux (EPN) d'enseignement supérieur en co-tutelle MC-MESR, des établissements publics (EP) et services à compétence nationale (SCN) ayant une vocation de recherche en lien avec la culture, auxquels on peut ajouter les unités mixtes de recherche (UMR) en co-tutelle avec le CNRS. La liste complète des unités prises en compte figure à l'annexe I).

Les unités du périmètre ont des activités de recherche dans des domaines variés qui se répartissent dans les deux grands champs de la recherche, les sciences de la vie et de la matière d'une part et les sciences humaines et sociales d'autre part. Les productions de la recherche, et notamment les publications, varient sensiblement entre ces deux grands champs. La recherche dans le domaine de la bibliométrie a étudié ces variations interdisciplinaires et le fait qu'en SHS, les articles scientifiques et les actes de colloques sont relativement moins fréquents qu'en sciences de la vie et de la matière (Hicks 2004 ; Nederhof 2006 ; Huang et al. 2008). Or les bases de données bibliométriques internationales, à l'instar du Web of Sciences ou de Scopus, recensent majoritairement des articles publiés dans des revues scientifiques et des actes de colloques. En particulier, et au moins jusqu'au milieu des années 2010, elles recensent peu d'ouvrages et de chapitres d'ouvrages (Norris et al. 2007 ; Sivertsen et al. 2012 ; Chi 2014a ; 2014b ; Faggiolani et al. 2018). Ces bases de données favorisent également les publications écrites en anglais (Waltman 2016 ; Mongeon et al. 2016).

Les indicateurs calculés à partir de ces bases de données peuvent présenter des biais s'agissant des publications des domaines des SHS dans la mesure où les ouvrages par exemple présenteraient des caractéristiques différentes des articles de revue en termes de thématiques ou d'impact (Hicks 1999 ; Norris et al. 2007 ; Waltman 2016). La meilleure couverture des bases de données pour les publications en anglais entraîne par ailleurs une sous-représentation des productions de pays non-anglophones. Les comparaisons de ce point de vue peuvent chercher à établir des référentiels de pays ou des institutions non-anglophones (OST 2018).

Afin de recenser et de caractériser la production des entités du périmètre institutionnel MC, l'analyse a été développée à partir de trois sources de données. Ces trois sources ont été choisies pour leur complémentarité : une base bibliométrique internationale permettant d'analyser les citations et les collaborations internationales, l'archive nationale HAL où les chercheurs et institutions français peuvent déposer des notices ou des documents complets et les listes des productions que les entités de recherche fournissent dans le cadre du processus d'évaluation par le Hcéres. Chacune de ces sources présente des avantages et des inconvénients pour recenser et caractériser la production des unités de recherche. Les indicateurs qui peuvent être produits varient donc d'une source à l'autre.

La première partie présente plus en détail les sources des données et la méthode suivie pour développer l'analyse des publications et plus largement des productions des unités du périmètre institutionnel MC. Les trois parties suivantes présentent successivement les résultats obtenus sur les trois corpus constitués : MC-WoS, MC-HAL et MC-Hcéres. La cinquième partie établit une comparaison systématique des résultats sur les corpus pour les indicateurs pour lesquels c'est possible. La dernière partie conclut et suggère des pistes d'approfondissement pour caractériser plus précisément les productions du périmètre institutionnel MC.

## DONNEES ET METHODE

### 1 / Sources des données sur la production des structures MC

#### Corpus issus du Web of Science et de HAL

La première source utilisée est la base de publications de l'OST, qui est une version enrichie de la base Web of Science (WoS) de Clarivate Analytics. L'OST enrichit la version complète du WoS, notamment avec des repérages précis des publications des institutions de l'enseignement supérieur et de la recherche françaises. La base utilisée pour cette étude a été actualisée en février 2017. L'année de publication la plus récente disponible à cette date est 2016 pour laquelle les données sont complètes à 80 -85%.

Le repérage des publications a été effectué sur l'ensemble de la base WoS (SCI-Science Citation Index Expanded, SSCI-Social Sciences Citation Index, A&HCI-Arts & Humanities Citation Index, CPCI-Conference Proceedings Citation Index (S et SSH)), quel que soit le type de documents.

HAL, archive ouverte, créée en 2001 par le CNRS, permet aux chercheurs de déposer leurs articles et manuscrits. Cette deuxième base de données est uniquement bibliographique (sans décompte des citations) et destinée au dépôt et à la diffusion de travaux, publiés ou non, y compris des thèses, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés. L'accès aux données est libre<sup>3</sup>, contrairement à l'accès au WoS qui nécessite un abonnement.

Cette base de données a été intégrée dans l'étude dans la mesure où c'est une base française plus représentative des productions nationales et/ou en langue française que le WoS, notamment pour les Sciences Humaines et Sociales.

Choisie par les établissements français de recherche, les universités et les grandes écoles au terme d'un protocole d'accord signé en juin 2006, puis par une convention de 2013, HAL propose un ensemble de services et d'outils aux institutions et aux chercheurs :

- transfert automatique des documents vers une archive ouverte internationale telle qu'ArXiv<sup>4</sup> ou Pubmed Central<sup>5</sup> lorsqu'ils appartiennent aux disciplines concernées ;
- création d'interfaces paramétrables permettant à des institutions ou des communautés scientifiques de créer leur propre environnement ;
- constitution de collections grâce à des « tampons » permettant d'authentifier la production d'un laboratoire, d'une équipe, les articles d'un journal... ;
- extraction automatique de listes de publications pour des chercheurs ou des laboratoires sous des formats divers.

Le dépôt des données par les chercheurs est volontaire : la complétude des données dépend donc des pratiques en matière de dépôt et des aides mises en place par les institutions pour faciliter les dépôts.

Le WoS et HAL permettent la sélection des publications afin de créer les corpus MC, à partir des :

- noms de structures auxquelles sont affiliés les auteurs ;
- noms des auteurs.

Les deux approches ont d'abord été combinées<sup>6</sup>. Cependant, l'approche par les noms d'auteurs a été abandonnée car les homonymies sont trop fréquentes et auraient nécessité un travail important pour ne garder que les publications du périmètre MC.

#### Création du corpus MC WoS : utilisation des intitulés des structures de rattachement (affiliation)

Pour créer le corpus, les différentes graphies des libellés de structures liées au MC (adresses) dans les signatures des publications ont été recherchées. À titre d'illustration, pour l'INRAP, les graphies suivantes ont été repérées :

- FRENCH NATL INST PREVENT ARCHAEOLOGICAL
- INRAP
- INST NATL RECH ARCHEOL PREVENT

<sup>3</sup> Source : [https://fr.wikipedia.org/wiki/Hyper\\_articles\\_en\\_ligne](https://fr.wikipedia.org/wiki/Hyper_articles_en_ligne)

<sup>4</sup> <https://arxiv.org/>

<sup>5</sup> <https://www.ncbi.nlm.nih.gov/pubmed/>

<sup>6</sup> voir CR réunion intermédiaire du 22 septembre 2017

- INST NATL RECH ARCHEOLOG PREVENT
- INST NATL RECHERCHES ARCHEOL PREVENT

Il avait été convenu, lors de la réunion du 22 septembre 2017, que le MC fiabiliserait le fichier des adresses afin de fournir des indications complémentaires sur les graphies utilisées pour identifier la structure d'affiliation de l'auteur dans les publications. En fait, cette fiabilisation a été prise en charge par l'OST ; une validation et un enrichissement du fichier des graphies par le MC constitueraient un complément utile pour renforcer la qualité du corpus obtenu.

### Création du corpus MC-HAL

Dans HAL, la création du corpus a suivi le même principe, à savoir l'utilisation des graphies différentes et/ou successives pour une même structure sous tutelle ou co-tutelle du MC. Le référentiel de structures AURÉHAL<sup>7</sup> propose, pour une même structure, différents enregistrements : pour chacune, toutes les formes ont été prises en compte afin de créer le corpus. Pour chaque structure de rattachement, toutes les publications ont été prises en compte, quels que soient l'auteur et son rattachement institutionnel. Ce périmètre prend donc en compte l'ensemble de la recherche du MC (voir Accord cadre MC-CNRS 2016-2020).

### Exemple : structures trouvées dans AURÉHAL pour le C2RMF

N° AURÉHAL	Libellé	Acronyme	Adresse
510	Laboratoire de Recherche et de Restauration des Musées de France	C2RMF	Palais du Louvre 14 quai François Mitterrand 75001 PARIS
110618	Laboratoire du Centre de Recherche et de Restauration des Musées de France	LC2RMF	Quai François Mitterrand 75001 Paris
150592	Centre de Recherche et de Restauration des Musées de France	C2RMF	UMR 171 PALAIS DU LOUVRE PORTE DES LIONS QUAI FRANCOIS MITTERRAND 75001 PARIS
241878	Centre de Recherche et de Restauration des Musées de France	C2RMF	Palais du Louvre - Porte des Lions 14 Quai François Mitterrand 75001 Paris
419539	Centre de Recherche et de Restauration des Musées de France	C2RMF	14 Quai François Mitterrand 75001 Paris

### Les données issues des dossiers d'autoévaluation du Hcéres

Les structures de recherche françaises fournissent, dans le cadre des évaluations menées par le Hcéres, un rapport d'autoévaluation comprenant notamment la liste des productions réalisées durant la période écoulée depuis la précédente évaluation. Ces évaluations sont organisées par vagues annuelles regroupant les établissements et/ou structures de recherche suivant une logique géographique. Compte tenu de la répartition géographique des structures relevant du MC, les dossiers pris en compte relèvent des 5 vagues successives d'évaluation entre 2007 et 2016. Les informations fournies par chaque structure ne concernent donc que les années précédant cette évaluation et ne couvrent que 5 années glissantes (et non continues) selon les vagues ; les dates de production ne permettront pas de donner d'indication d'évolution.

Par ailleurs, toutes les structures à prendre en compte dans le cadre de cette étude ne participent pas à une évaluation par le Hcéres (voir liste en Annexe I). Les structures suivantes ont été prises en compte :

- Ensemble des UMR ;
- EPN (Musée du Quai Branly) ;
- UR des ENSA.

Sur les 155 structures recensées dans le cadre de ce rapport, 50 ont fourni un rapport d'autoévaluation duquel ont été extraites les listes de production.

Les données utilisées dans cette étude proviennent des listes de publications et autres productions fournies par les structures lors de la constitution de leur dossier d'évaluation pour le contrat quinquennal d'établissement. Ces listes, directement fournies par les structures sont a priori exhaustives.

Toutes les productions ont été retenues, sans tenir compte d'un éventuel statut du producteur (membre statutaire, doctorant...) et un type de production leur a été attribué. Il n'y a eu aucun retraitement des données. Les mentions imprécises ou cumulatives (« ...notes des textes liminaires de 6 pièces... », « ...35 notices scientifiques... ») ont été prises in extenso. Il n'y a pas de dédoublonnage en raison même de l'hétérogénéité

<sup>7</sup> <https://aurehal.archives-ouvertes.fr/>

des données source. Enfin, les productions déclarées sont reprises telles quelles, sans vérification de leur référencement dans un corpus bibliographique de dépouillement de documents publiés (base bibliographique publique ou commerciale).

Les unités de recherche appartenant au périmètre du MC ayant été évaluées entre 2012 et 2016, leur production scientifique porte sur les années allant de 2007, pour les unités évaluées lors de la vague D, à 2016 pour les unités évaluées lors de la vague C. Cette période a donc été retenue pour caractériser la recherche sous tutelle du MC.

## 2 / Définitions et calcul d'indicateurs

Dans le cadre de cette étude, le terme de « **publication** » renvoie aux différentes formes de publication scientifique : les articles dans des revues à comité de lecture, les contributions à des colloques ayant donné lieu à publication d'actes notamment. Il permet de les distinguer du reste de la **production** des unités, terme plus général. La production dans un sens large comprend à la fois la production scientifique et la production d'ouvrages de vulgarisation par exemple.

Dans le WoS, l'indexation disciplinaire est réalisée au niveau de la revue et non au niveau de la publication ; ainsi, un article « hérite » t-il de la, ou des, **discipline(s)** de la revue dans laquelle il est publié. Dans HAL, ce sont des **domaines de recherche** qui sont utilisés lors de l'indexation réalisée par les déposants : cette indexation concerne le produit déposé.

Les données sont analysées en **compte de présence** : dans une logique de « participation » à la production scientifique, chacune des publications dans lesquelles l'adresse de la structure apparaît est comptabilisée « 1 » quel que soit le nombre de contributeurs à la publication. Cette logique est reproduite au niveau des disciplines. Lorsqu'une publication relève de deux disciplines, elle est comptabilisée « 1 » dans chacune.

**La spécialisation** scientifique dans une discipline signifie qu'un acteur (pays ou institution) a une part relativement plus importante de sa production scientifique dans cette discipline que la zone de référence (le monde en général, la France dans ce rapport). Dans ce rapport, l'indice de spécialisation rapporte la part d'une discipline dans le total des publications d'un acteur à ce même ratio pour la France. Par construction, la valeur neutre de l'indice de spécialisation est 1.

**L'impact scientifique** ou académique est mesuré à partir des citations reçues par les publications. La citation permet aux auteurs de renvoyer aux travaux antérieurs qui ont porté sur le même sujet, dans le but de reconnaître l'antériorité, comparer et/ou critiquer. La citation est considérée comme un indice permettant de présumer de l'influence académique des travaux de recherche. **L'indice d'impact** est le nombre de citations par les publications d'un acteur, rapporté au nombre de citations par publication d'une zone de référence. Par construction, l'indice d'impact est 1 pour le monde ou la zone de référence. La **fenêtre de citation** représente le délai, incluant l'année de publication, pris en compte pour comptabiliser le nombre de citations reçues par une publication. Dans ce rapport, une fenêtre de 3 ans est utilisée.

Le corpus du MC a été analysé en faisant appel à une **modélisation thématique**. La méthode s'appuie sur la fréquence de cooccurrence des mots utilisés dans le titre et le résumé des publications pour révéler les principaux thèmes présents dans ces publications. Elle permet de visualiser le corpus du MC en fonction des thèmes ainsi identifiés : compte tenu des volumes, le paramétrage du logiciel a été défini autour de 15 thèmes. L'analyse de ces derniers, à partir des mots les plus fréquents ou les plus pertinents, ainsi que la lecture des titres et résumés des 100 articles les plus représentatifs de chaque thème, a permis à l'OST, non spécialiste du domaine, de donner un titre aux thèmes puis de dessiner des clusters.

Une **publication mono-adresse** ne comporte l'adresse que d'une seule institution. Une **co-publication** est une publication co-signée par au moins deux institutions différentes, publiques et/ou privées, domiciliées dans le même pays et/ou dans des pays différents. Le **réseau de collaborations** du périmètre institutionnel MC est approché par une analyse des co-publications entre une structure MC et une autre structure, en France ou à l'étranger. Le logiciel VOSviewer<sup>8</sup> qui génère des cartes sur la base de réseaux de données bibliométriques a été utilisé. La carte est construite à partir des co-publications entre institutions ou pays. Lorsqu'une publication est co-signée par des auteurs affiliés à deux ou plusieurs institutions, cette co-publication définit de fait un lien entre les institutions impliquées. La largeur du lien mesure l'intensité de collaboration entre les institutions. Deux institutions appartiennent à un même groupe lorsque la distance obtenue en combinant les liens qui les unissent est faible. La taille des bulles est proportionnelle au nombre de publications.

---

<sup>8</sup> Logiciel libre développé par le CWTS de l'université de Leiden.

## ANALYSE DU CORPUS MC-WOS

### 1 / Production du MC recensée dans le WoS

À partir des intitulés des structures de recherche en tutelle ou co-tutelle du MC (voir liste en Annexe I), un corpus de publications sur la période 2007-16 a été créé à partir des données disponibles dans la base de publications scientifiques enrichie par l'OST.

**Tableau 1 : Évolution du nombre de publications WoS-MC, 2007-16**

	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016*	2007-16*
<b>Nombre de publications</b>	201	243	273	331	334	381	418	443	493	477	3 594

Source : Clarivate Analytics, traitements OST

\* Année 2016 incomplète

Le nombre de publications du corpus MC recensées dans le WoS a nettement progressé entre 2007 et 2016, passant de 201 publications en 2007 à 493 en 2015 (dernière année complète). Soit une progression de plus de 145%.

Cette progression plus importante que la moyenne nationale (51% pour les publications en SHS) peut être expliquée par un effort des chercheurs du MC à publier en anglais dans des revues internationales à comité de lecture mais également par la politique menée par *Clarivate Analytics* pour intégrer de nouvelles revues en SHS.

Les publications recouvrent plusieurs types de documents dont la définition est disponible en ligne. Afin de se rapprocher de ce qui est présent dans les autres corpus développés pour ce rapport, une typologie, composée de 4 catégories (voir descriptif en Annexe II) a été utilisée.

Cependant, le tableau suivant ne comporte que 3 catégories car le corpus MC ne comporte pas d'ouvrages : ces derniers sont peu présents dans le WoS (et seulement en anglais) et représentent seulement 10 % des publications en SHS.

Dans le WoS, l'indexation des types de documents peut être multivaluée, notamment pour les articles : un article peut ainsi également être un *Proceedings Paper* (248 occurrences) ou un Book chapter (1 occurrence). Dans le corpus présent, seule l'indication « article » a été retenue. Les publications classées dans la catégorie Article ont été soumis à un processus de revue par les pairs dans le cadre d'une revue ou de la publication d'actes d'un colloque.

**Tableau 2 : Nombre de documents par type dans le corpus MC-WoS**

		2007	2008	2009	2010	2011	2012	2013	2014	2015	2016*	2007-16*	
<b>Article</b>	Article	171	215	231	284	290	322	363	394	424	419	3 113	
	Editorial Material	10	4	12	20	9	17	7	14	20	19	132	
	Review	6	10	13	6	8	5	6	6	10	7	77	
	Letter			2		2	3	2	1	1	3	14	
	Correction, Addition		1		1				2	1	2	3	10
	Reprint						1						1
												<b>3 347</b>	
<b>Congrès</b>	Meeting Abstract	6		4	11	8	13	17	11	7	4	81	
	Book Review	5	8	10	8	14	16	13	10	26	20	130	
<b>Autre production</b>	Art Exhibit Review	2	2	1		1	2	2	3	1	1	15	
	New Item					1	2	4	2	1	1	11	
	Biographical-Item	1	2		1	1		1	1	1		8	
	Bibliography		1									1	
	Poetry							1				1	
												<b>166</b>	
<b>Total annuel</b>		<b>201</b>	<b>243</b>	<b>273</b>	<b>331</b>	<b>334</b>	<b>381</b>	<b>418</b>	<b>443</b>	<b>493</b>	<b>477</b>	<b>3 594</b>	

Source : Clarivate Analytics, traitements OST

\* Année 2016 incomplète

Le corpus des publications du MC est constitué d'un large panel de types de documents mais les publications recensées dans le WoS sont composées d'articles à 93 % (tableau 3). Il est à noter que le profil des publications françaises, toutes disciplines confondues sur la même période, est composé à 66 % d'articles publiés dans des revues à comité de lecture. La part des meetings abstract (résumé des présentations faites lors de congrès ≠ acte de congrès publié) est faible tout comme celle des « autres productions ».

**Tableau 3 : Distribution des documents par type, 2007-16**

	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016*	2007-16*
<b>Article</b>	93,0%	94,7%	94,5%	94,0%	92,5%	91,3%	90,9%	93,9%	92,7%	94,5%	93,1%
<b>Ouvrage</b>	-	-	-	-	-	-	-	-	-	-	-
<b>Meeting abstract/ Congrès</b>	3,0%	0,0%	1,5%	3,3%	2,4%	3,4%	4,1%	2,5%	1,4%	0,8%	2,3%
<b>Autre production</b>	4,0%	5,3%	4,0%	2,7%	5,1%	5,2%	5,0%	3,6%	5,9%	4,6%	4,6%

Source : Clarivate Analytics, traitements OST

\* Année 2016 incomplète

## 2 / Langues des publications

Le WoS indexe des revues dans lesquelles le titre et le résumé des publications doivent être fournis en anglais. Cependant, la publication elle-même peut avoir été rédigée dans une autre langue.

**Tableau 4 : Nombre de documents du corpus MC-WoS selon la langue, 2007-16**

	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016*	2007-16*
<b>Anglais</b>	131	161	166	235	235	266	317	347	391	392	2 641
<b>Français</b>	66	76	101	89	89	111	96	87	97	81	893
<b>Espagnol</b>	2	3	6	1	7	3	3	4	4	2	35
<b>Allemand</b>	3	1		4	2	1	2	4		2	19
<b>Autres</b>		1		2	1			1		1	6
<b>Total</b>											3 594

Source : Clarivate Analytics, traitements OST

\* Année 2016 incomplète

Alors que les publications françaises, toutes disciplines confondues sur la même période, sont publiées à 91 % en langue anglaise et à 8 % en langue française, celles du MC sont davantage en langue française. Cette proportion baisse sensiblement sur la période, passant de 33 % à moins de 20 %.

**Graphique 1 : Évolution de la part des publications du MC-WoS en français et en anglais, 2007-16**



Source : Clarivate Analytics, traitements OST

\* Année 2016 incomplète

Pour la caractérisation de la recherche culturelle, la combinaison des supports de publication (type de document) avec la langue de publication apporte des informations complémentaires. En effet, les différents types de production (article, congrès, ouvrage ou autre production) ne seront pas systématiquement rédigés en anglais étant donné que les publics varient selon deux axes, d'une part, la communauté nationale versus une audience internationale et, d'autre part, la recherche versus la diffusion des connaissances.

**Tableau 5 : Répartition annuelle des langues par type de document MC-WoS**

		2007	2008	2009	2010	2011	2012	2013	2014	2015	2016*
<b>Article</b>	Article Ang	63%	67%	60%	71%	70%	72%	76%	78%	81%	84%
	Article Fra	34%	31%	38%	27%	27%	27%	23%	19%	18%	15%
<b>Meeting abstract</b>	Congrès Ang	100%	-	100%	100%	100%	100%	100%	100%	100%	100%
<b>Autre prod.</b>	Autre prod. Ang	75%	54%	73%	33%	59%	20%	57%	63%	48%	45%
	Autre prod. Fra	25%	31%	27%	67%	35%	80%	38%	38%	52%	55%

Source : Clarivate Analytics, traitements OST

\* Année 2016 incomplète

Les articles des structures MC sont publiés de plus en plus en anglais. Les communications dans un congrès sont toutes en anglais.

### 3 / Domaines de recherche et thématiques du corpus MC-WoS

La part du corpus MC-WoS dans les publications françaises des domaines de recherche où le MC compte au moins une publication dans le WoS est de 0,3% en 2015. Le calcul de la part du MC sur cet ensemble implique que l'on compare le corpus MC au total des publications de la France toutes disciplines confondues. Cela introduit un biais car les deux tiers des publications du MC sont concentrés dans 10 domaines de recherche. En conséquence, le périmètre d'analyse, à des fins de comparaisons, a été restreint aux 10 premiers domaines de recherche.

La part du corpus MC-WoS 10 dans les publications françaises est passée de 2,8 à 5,1 % entre 2007 et 2016. Cette part est de 15,7 % en sciences humaines et sociales (SHS) en 2016 et de 3,6 % en sciences de la matière et de la vie (SMV), comme le montre le graphique 3. Malgré le faible nombre de publications, ce corpus représente un poids non négligeable dans le paysage de la production scientifique française, y compris lorsqu'on considère le domaine des SMV. Au sein des SHS, qui correspondent plus directement aux missions des structures du MC, les valeurs sont plus élevées.

**Graphique 2 : Part du corpus MC-WoS 10 dans les publications françaises par grand domaine scientifique**



Source : Clarivate Analytics, traitements OST

\* Année 2016 incomplète

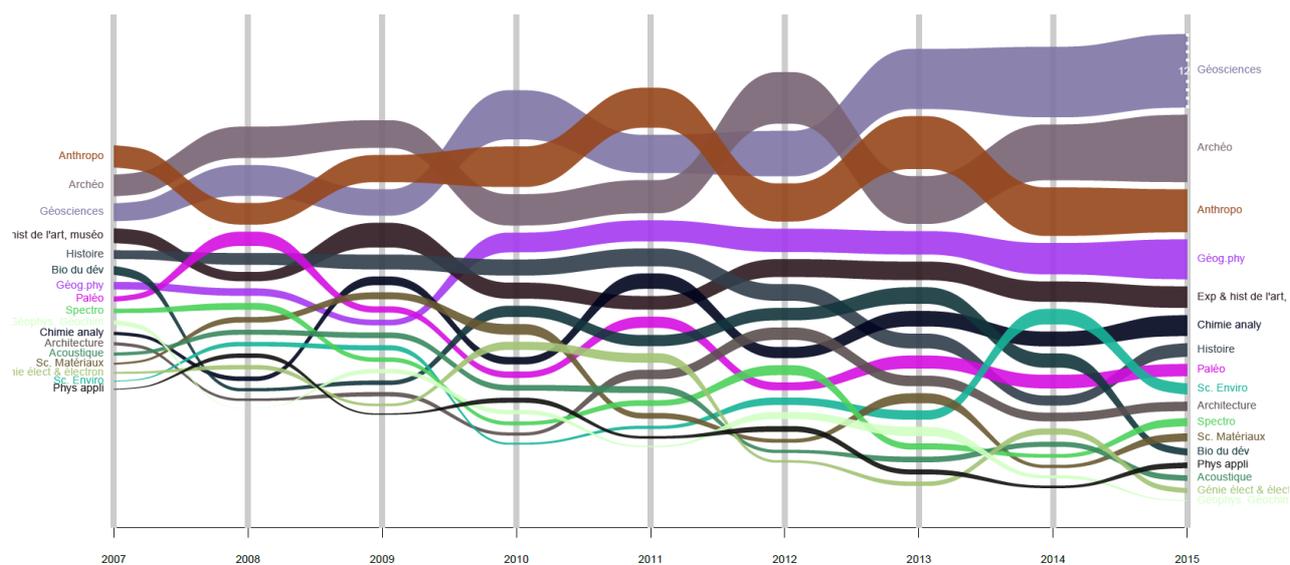
Les domaines de recherche du corpus MC-WoS 10 sont précisés par le tableau 6 et le graphique 3.

**Tableau 6 : Nombre de documents par discipline dans le MC-WoS 10, 2007-16**

	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016*	2007-16*	Part**	
SHS	Archéologie	35	51	45	51	54	84	78	91	110	60	<b>659</b>	<b>18,3%</b>
	Anthropologie	36	35	44	66	65	62	86	79	70	61	<b>604</b>	<b>16,8%</b>
	Géogr. physique	12	12	9	32	34	38	38	51	65	108	<b>399</b>	<b>11,1%</b>
	Expression artistique, hist. de l'art, muséographie	24	15	40	26	22	29	30	33	35	22	<b>276</b>	<b>7,7%</b>
	Histoire	14	19	24	26	29	27	24	16	22	23	<b>224</b>	<b>6,2%</b>
	Paléontologie	8	23	10	10	18	13	21	22	20	15	<b>160</b>	<b>4,5%</b>
	Architecture	5	4	7	5	15	20	17	14	15	16	<b>118</b>	<b>3,3%</b>
SMV	Géosciences multidisciplinaires	29	50	43	80	62	74	98	115	120	141	<b>812</b>	<b>22,6%</b>
	Chimie analytique	5	7	14	12	25	18	25	24	34	24	<b>188</b>	<b>5,2%</b>
	Sciences de l'environnement	1	7	8	3	5	12	15	25	18	10	<b>104</b>	<b>2,9%</b>
<b>Total MC-WoS 10</b>	<b>118</b>	<b>154</b>	<b>186</b>	<b>212</b>	<b>232</b>	<b>245</b>	<b>288</b>	<b>299</b>	<b>320</b>	<b>300</b>	<b>2 354</b>	<b>100%</b>	

\* Année 2016 incomplète ; \*\* le calcul tient compte du fait que les publications sont en compte de présence, donc la somme des disciplines est supérieure au nombre de documents dans le corpus qui correspond à la ligne total.  
Source : Clarivate Analytics, traitements OST

**Graphique 3 : Nombre de publications dans le MC-WoS 10 par domaine de recherche, 2007-15**

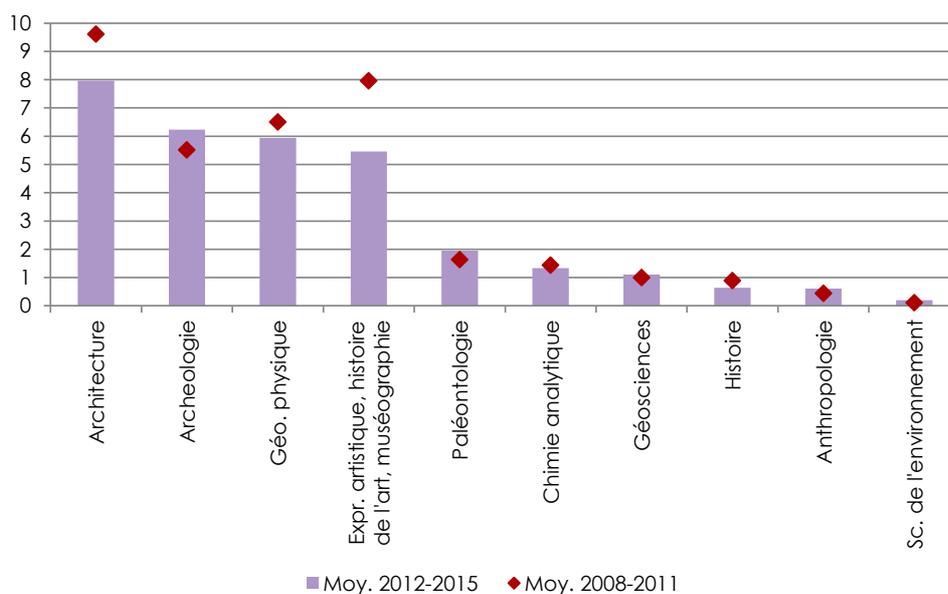


Source : Clarivate Analytics, traitements OST

Le tableau 6 et le graphique 3 montrent que trois disciplines se distinguent nettement par leur volume de publication. Les géosciences multidisciplinaires, relevant des sciences de la terre, sont le premier domaine de publication. Il s'agit d'approches interdisciplinaires pour aborder l'étude de la terre et des autres planètes qui, d'après la nomenclature du WoS, portent sur la géologie, la géochimie/géophysique, l'hydrologie, la paléontologie, l'océanographie, la météorologie, la minéralogie, la géographie et l'énergie. Viennent ensuite des disciplines, relevant des sciences humaines, de l'anthropologie et de l'archéologie.

Associés plus intuitivement au MC, la paléontologie qui relève des sciences humaines, la géologie ou la géographie participent à la composition de la discipline des géosciences au sein du WoS, explicitant ainsi la présence des géosciences dans ce corpus. L'analyse des revues de géosciences multidisciplinaires où l'on retrouve des publications du corpus MC-WoS montre que les cinq premières revues portent sur le quaternaire, l'holocène et l'archéologie. 64 de ces revues en géosciences sont communes avec la géographie physique, 50 avec la paléontologie et 345 avec l'archéologie. Enfin, l'analyse des thématiques de ces publications montre qu'elles portent principalement sur les colonies humaines durant l'âge de fer, l'impact humain durant l'holocène et le paléo-environnement.

**Graphique 4 : Spécialisation par rapport à la France, par domaine de recherche du corpus MC-WoS 10**



Source : Clarivate Analytics, traitements OST

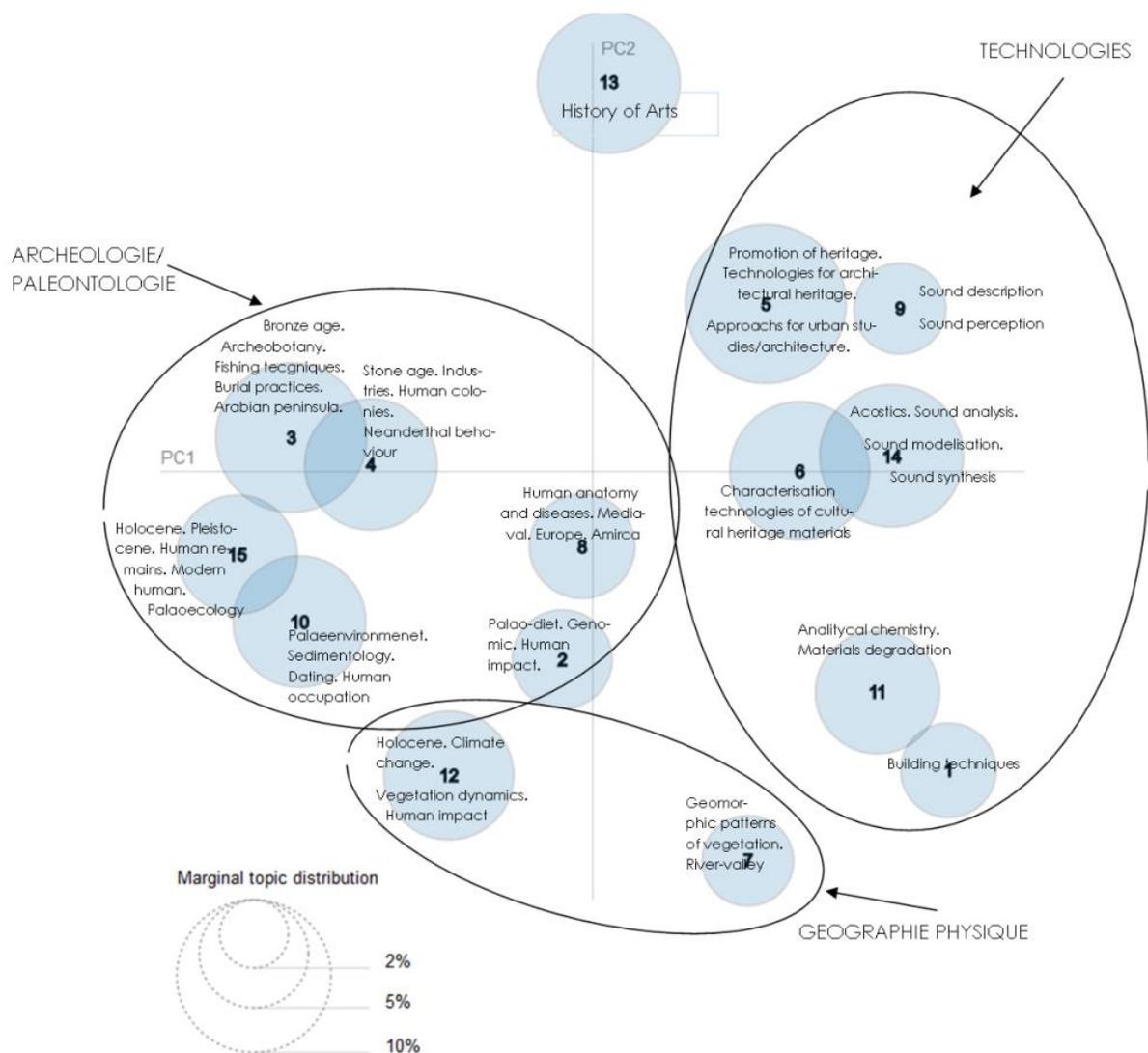
Le groupe des disciplines de plus forte spécialisation au sein du corpus est composé de l'architecture, de l'archéologie, de la géographie physique et de l'expression artistique/ histoire de l'art/ muséographie, même si leurs indices tendent à baisser sur la 2<sup>ème</sup> période 2012-2015.

Le corpus est beaucoup moins spécialisé en paléontologie et en chimie analytique. La part des géosciences multidisciplinaire dans le corpus est équivalente à celle de l'ensemble des publications de la France (valeur neutre de l'indice à 1). Enfin, le corpus n'est pas spécialisé en histoire, anthropologie et sciences de l'environnement.

La cartographie des thèmes (voir p. 9), basée sur 2 835 publications ayant un résumé en anglais (78% du corpus) fournit une présentation du contenu des articles complémentaire de la classification plus disciplinaire. La figure 1 identifie 4 groupes thématiques :

- « Archéologie/paléontologie » où les recherches portent sur les techniques, les industries développées par l'Homme ainsi que sur son évolution physique et comportementale sur une période allant de l'âge de pierre au moyen âge ;
- « Technologies et analyses du patrimoine » rassemble des publications diverses autour des technologies et des outils pour caractériser les matériaux et leur dégradation, des analyses du patrimoine, de l'architecture et de l'urbanisme, ou encore des analyses des sons et de leur perception ;
- « Géographie physique » porte sur le changement climatique, l'impact de l'homme et sur les dynamiques de la végétation dans certains milieux (rivières, vallées) ;
- « Histoire de l'art » (n°13) se dégage comme une thématique de recherche, tout en étant relativement proche du thème sur le patrimoine (n°5).

Figure 1 : Cartographie des thématiques de l'ensemble du corpus MC-WoS



Source : Clarivate Analytics, traitements OST

Remarque : les nombres ne servent qu'à repérer les différents topics, ils n'ont pas d'autre signification

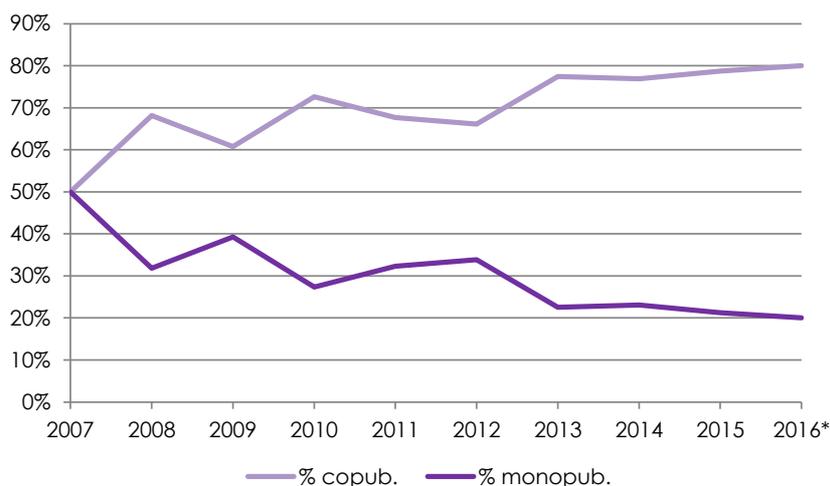
\* Année 2016 incomplète

## 4 / Co-publications et réseaux au sein du corpus MC-WoS

Sur le périmètre MC-WoS 10, les structures du MC publient de plus en plus en collaboration avec d'autres acteurs : 240 co-publications sur 300 en 2016 (soit 80 % des publications en co-publication), contre 59 sur 118 en 2007 (soit 50 % des publications en co-publication).

Cette évolution suit les tendances nationale et internationale : la part de co-publications entre institutions augmente et ceci, quelle que soit la discipline. Le rapport sur la position scientifique de la France dans le monde le souligne. Plus précisément, en ce qui concerne les sciences humaines (graphique 24 du rapport), 40 % des publications françaises en sciences humaines sont signées par une seule institution ; le corpus du MC est au-dessous de cette moyenne nationale avec 20 % de mono-adresses en 2016.

**Graphique 5 : Évolution du taux de co-publications du MC-WoS 10**

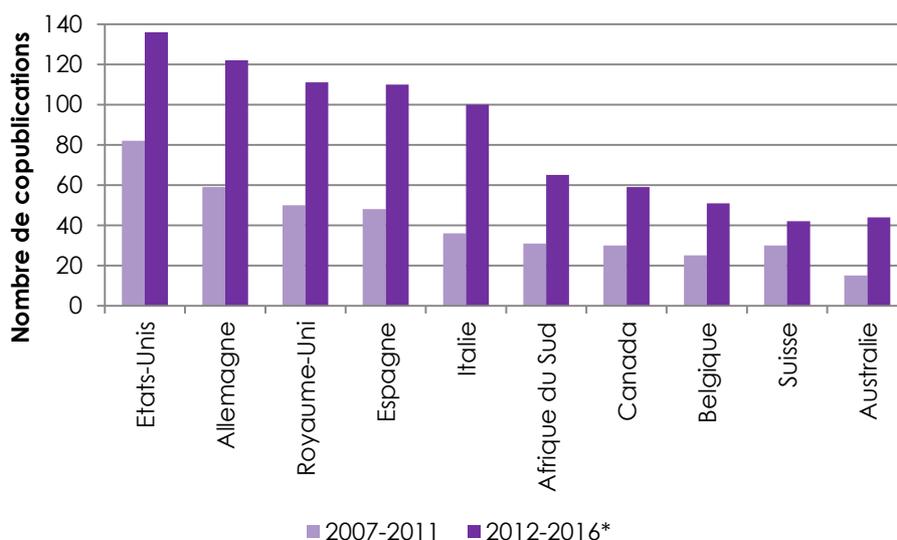


Source : Clarivate Analytics, traitements OST

\* Année 2016 incomplète

Comme pour l'ensemble des publications françaises (OST 2018), les premiers partenaires au sein du corpus MC-WoS 10 sont des institutions américaines suivies d'institutions de différents pays européens. Les co-publications internationales du corpus se font également avec l'Afrique du Sud et le Canada notamment (graphique 5).

**Graphique 6 : Les 10 premiers pays partenaires du MC.**



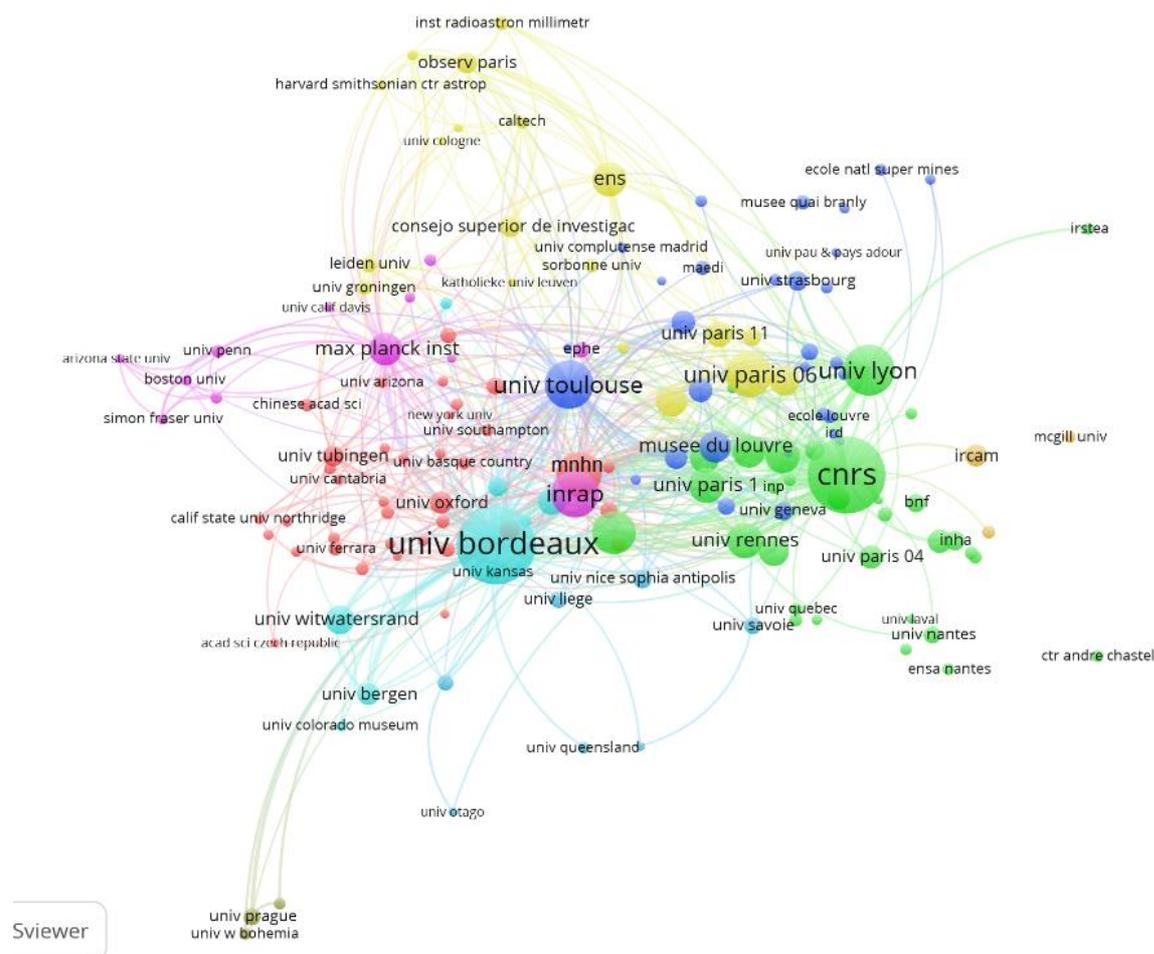
Source : Clarivate Analytics, traitements OST

\* Année 2016 incomplète

La carte des collaborations avec les autres pays, p.18, complète le tableau par une représentation graphique qui fait apparaître ces données sous forme de réseaux. Afin d'obtenir une carte lisible, n'ont été prises en compte que les co-publications du corpus ayant plus de 10 occurrences sur la période étudiée. Exemple : plus de 10 co-publications entre l'INRAP et l'Université de Lyon.

La largeur et la distance entre les différentes institutions indiquent l'intensité de la relation entre 2 structures. Dans ce réseau, les structures du MC sont approchées par l'une de leur tutelle. Les acteurs principaux sont l'Université de Bordeaux (bleu clair), le CNRS (vert), l'Université de Toulouse (bleu sombre), le MNHN (Muséum national d'histoire naturelle) (rouge), l'INRAP (rose), l'Université de Lyon (vert) et l'Université de Paris 06 (jaune).

**Figure 2 : Co-publications entre institutions du corpus MC-WoS\*, 2007 - 2016\*\***

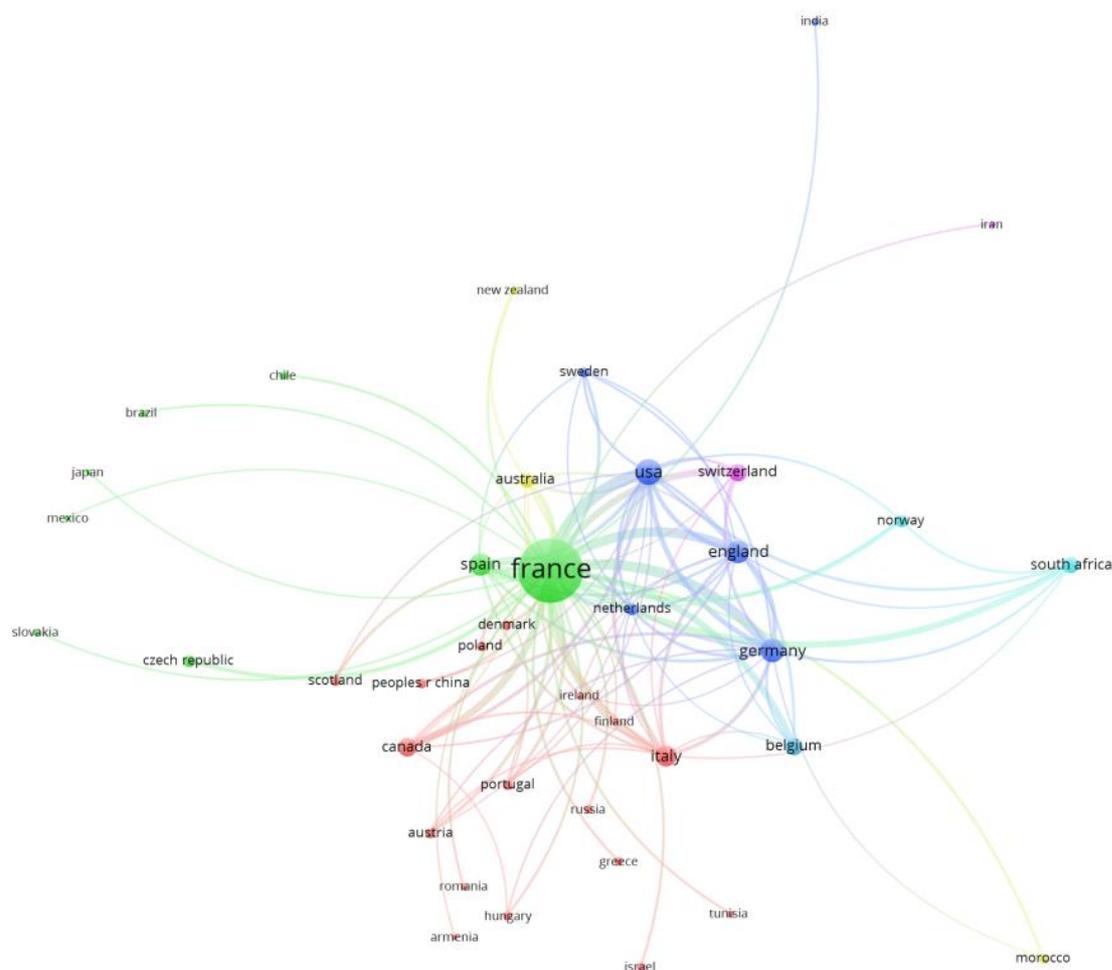


\* A partir de 10 co-publications entre 2007 et 2016      \*\* 2016 incomplète  
 Source : Clarivate Analytics, traitements OST avec VOSviewer

L'université de Bordeaux (réseau bleu clair) co-publie avec des universités en Afrique du Sud, en Europe et aux États-Unis. Le réseau du CNRS (réseau vert), et l'Université de Lyon dans son sillage, est composé notamment d'institutions françaises, à l'exception de l'Université du Québec. L'Université de Toulouse (réseau bleu sombre) a un réseau principalement franco-français avec le Musée du Louvre, le Synchrotron Soleil et quelques universités européennes. Le MNHN (réseau rouge) co-publie avec des universités européennes et dans une moindre mesure, avec la Chine et les États-Unis. Le réseau de l'INRAP (réseau rose) est international avec une forte collaboration avec le Max Planck Institute et dans une moindre mesure avec des universités nord-américaines. Enfin, le réseau de l'Université Paris 6 est franco-européen présentant quelques co-publications avec des universités américaines.

La cartographie des pays avec lesquels le MC co-publie montre une domination des co-publications franco-françaises suivie de co-publications avec les États-Unis, l'Angleterre, les pays européens (Allemagne, Espagne, Italie, Suisse et Belgique), le Canada et l'Afrique du Sud (graphique suivant).

Figure 3 : Réseaux de co-publications au sein du corpus MC-WoS\*, 2007 - 2016\*\*

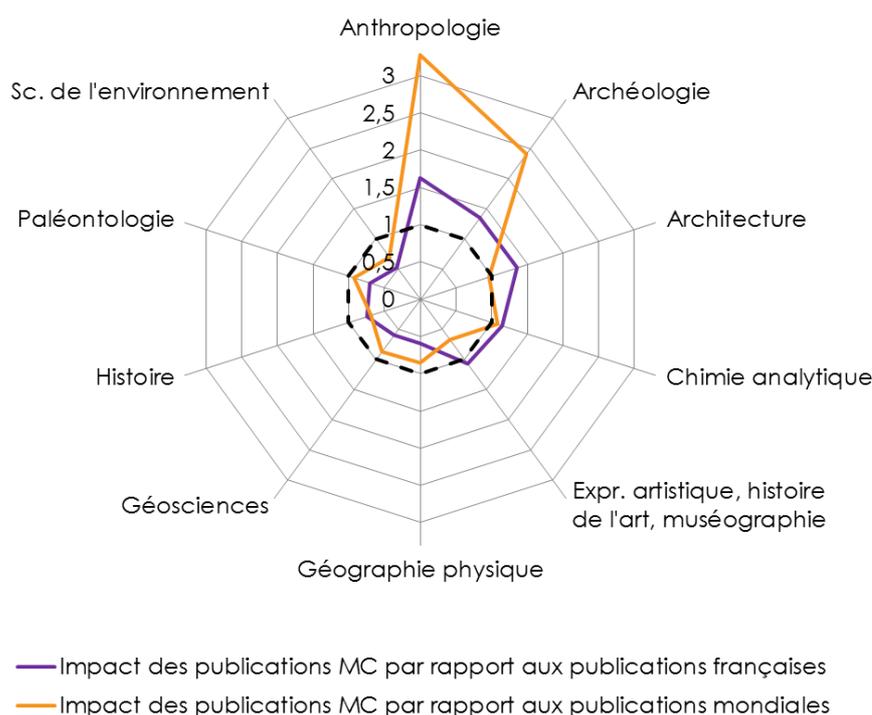


\* A partir de 10 co-publications entre 2007 et 2016      \*\* 2016 incomplète  
 Source : Clarivate Analytics, traitements OST avec VOSviewer

## 5 / Impact des publications du corpus MC-WoS

L'indice d'impact des publications du corpus MC-WoS 10 par rapport aux publications françaises, toutes disciplines confondues, est passé de 1,4 en 2007 à 1,7 en 2013. Les publications du corpus reçoivent donc, sur toute la période, plus de citations que la moyenne des publications françaises. Le graphique 7 précise les indices d'impact par domaine de recherche et compare le périmètre MC-WoS 10 par rapport aux publications de la France entière d'une part et du monde d'autre part. Les domaines de recherche ayant un indice supérieur à 1 ont un nombre de citations par publication supérieure à la moyenne nationale ou mondiale.

**Graphique 7 : Impact des publications du corpus MC-WoS 10 par domaine de recherche, 2007-13**



Source : Clarivate Analytics, traitements OST

L'anthropologie est le domaine de recherche dans lequel le corpus MC a les impacts les plus élevés : ses publications reçoivent 3,3 fois plus de citations par publication que la moyenne des publications mondiale et 1,6 fois plus que les publications françaises dans ce domaine. En archéologie le nombre de citations par publication est 2,4 fois plus élevé pour le corpus MC-WoS que pour le monde et 1,4 fois plus élevé que pour la France. Dans le domaine de l'architecture, les publications du MC reçoivent plus de citations que les publications françaises (1,4 fois) et un nombre équivalent à celui de l'ensemble des publications mondiales. L'écart est dû au fait que l'impact de l'ensemble des publications de la France en architecture a un indice d'impact de 0,7 (voir l'annexe IV).

En chimie analytique, les impacts par rapport au monde et à la France sont similaires (de l'ordre de 1,1).

En paléontologie, le corpus MC-WoS a un impact proche de la moyenne mondiale, mais inférieur à celui de l'ensemble des publications françaises.

## ANALYSE DU CORPUS MC-HAL

HAL étant une archive ouverte française, les productions déposées ne concernent que ce périmètre national, sans comparaison possible avec des données internationales.

Afin de caractériser le corpus du MC au sein de la production française déposée dans HAL, il est positionné non seulement par rapport à l'ensemble des dépôts toutes disciplines confondues, mais également par rapport aux dépôts dans le domaine des SHS (HAL dom. SHS). Ce périmètre spécifique, différent du portail HAL-SHS, est composé des productions pour lesquelles le déposant a utilisé au moins 1 fois un domaine lié aux SHS.

### 1 / Types de dépôt : archives et notices

Une base de données bibliographique comme HAL répertorie des articles, actes de congrès... décrits grâce aux métadonnées (auteurs et leur affiliation, mots-clés...). Ces dernières permettent aux utilisateurs de faire des recherches sur les documents. HAL a vocation à être une archive et doit non seulement référencer les documents primaires mais aussi les stocker.

HAL dispose d'un champ indiquant le type de dépôt effectué :

- Notice : fiche descriptive de la production.
- Archive : notice qui contient soit un fichier numérisé, soit un lien DOI vers la production en ligne.
- Annexe : notice qui contient un document attaché qui n'est pas une publication (document de présentation, fiche technique..).

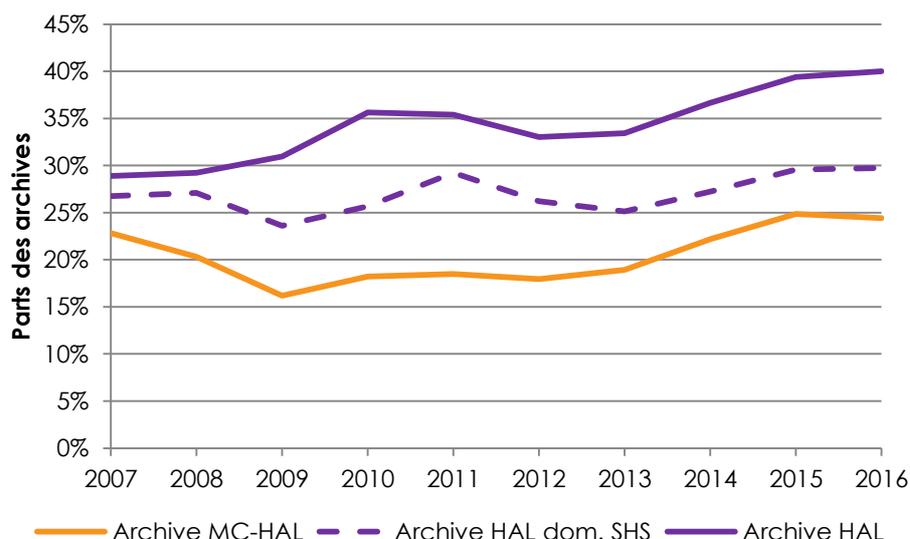
**Tableau 7 : Évolution des types de dépôt des structures MC (MC-HAL)**

	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2007-16
<b>Notice</b>	77,2%	79,6%	83,7%	81,6%	81,3%	81,8%	80,7%	77,4%	74,3%	74,6%	79,2%
<b>Archive</b>	22,8%	20,3%	16,2%	18,2%	18,5%	17,9%	18,9%	22,2%	24,9%	24,4%	20,4%
<b>Annexe</b>	0,0%	0,1%	0,1%	0,1%	0,2%	0,2%	0,3%	0,4%	0,9%	1,0%	0,4%

Source : HAL, traitements OST

À partir de 2009, les courbes des dépôts d'archives MC-HAL et HAL suivent une évolution similaire. Cependant, alors que dans l'ensemble de HAL, la proportion de fichiers déposés passe de 29 % à 40 % sur toute la période, pour le corpus MC-HAL les dépôts n'atteignent pas les 25 % des productions déposées.

**Graphique 8 : Part des archives dans les dépôts des corpus MC-HAL, HAL dom. SHS et HAL**



Source : HAL, traitements OST

Dans la suite du rapport, il n'est pas fait de distinction entre les types de dépôts ; ils sont tous pris en compte dans la mesure où l'objectif est ici de caractériser la production correspondante.

## 2 / Nombre de dépôts et types de productions du corpus MC-HAL

A la date de sa création, le 20/02/2018, le corpus MC-HAL comporte 29 244 produits pour la période 2007 - 2016. Les dépôts annuels sont passés de 2 260 en 2007 à 3 088 en 2015. Le terme « production » est utilisé dans la mesure où HAL contient non seulement des publications, mais également des rapports, des thèses, des cours... Le nombre de dépôts des structures MC dans HAL augmente jusqu'en 2011, tend à se stabiliser puis à se tasser (tableau 8).

**Tableau 8 : Nombre de dépôts MC dans HAL**

	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2007-16
Nombre de productions MC	2 260	2 535	2 699	2 788	2 873	3 317	3 365	3 355	3 088	2 944	29 224
Parts dans HAL-SHS	9,6%	9,8%	9,6%	9,1%	8,8%	9,1%	9,3%	9,6%	9,2%	9,1%	9,3%
Parts dans HAL	2,8%	2,9%	3,0%	2,8%	2,9%	3,2%	3,2%	3,3%	3,1%	3,0%	3,0%

Source : HAL, traitements OST

La part du corpus MC représente 3 % des productions déposées dans HAL et 9 % des productions du périmètre HAL dom. SHS.

Les dépôts des structures sous tutelle MC suivent la tendance des dépôts toutes institutions confondues dans HAL. En effet, les parts de production du MC dans le corpus général, comme dans le corpus HAL dom. SHS, ne marquent pas d'évolution notable sur la période.

Afin d'étudier l'évolution des productions, les 25 différents types de productions disponibles dans HAL ont été agrégés en 9 grands types (tableau 9). Le corpus MC-HAL comporte des thèses (533) et des mémoires d'étudiants (272). Il comporte aussi d'autres produits non publiés. Les rapports de fouilles appartiennent à la catégorie « Autres, non publié », a priori dans la sous-catégorie rapport (855). La proportion de cette catégorie Autres augmente sur la fin de la période, passant de 3,2 % en 2007 à 8,6% en 2015. L'ensemble des documents non publiés comporte en outre les thèses et les communications dans des séminaires.

Dans la base WoS, une communication dans un congrès, si elle est publiée dans une revue à comité de lecture, prend le statut d'article. Le tableau 9 fait apparaître ces communications qui dans MC-HAL signalent un support de publication et dont le statut peut être revu : la répartition de la production se trouve alors modifiée comme l'indique le graphique 9 (Type de prod. Revu). Ce dernier distingue les publications et permet de voir que les documents qui ne correspondent pas à des publications représentent 30 % du total.

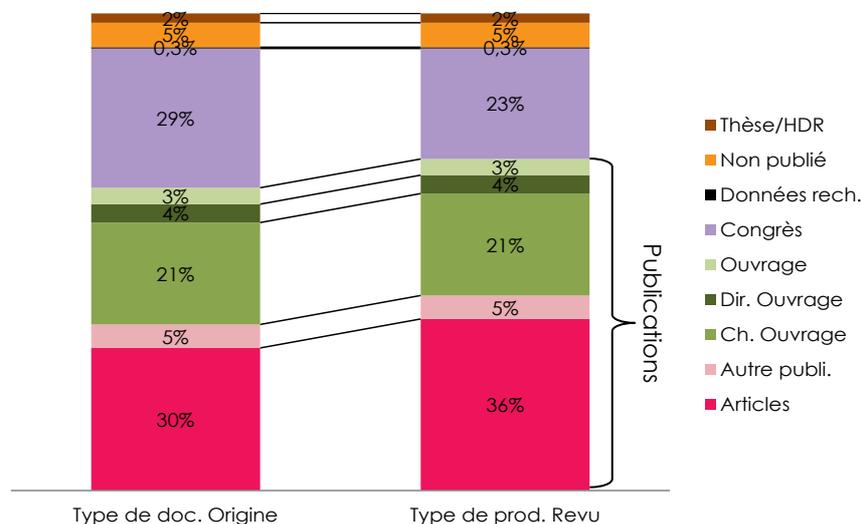
**Tableau 9 : Volumes par type de production dans HAL-MC**

	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2007-16	
<b>Article</b>	Article de revue	608	764	740	880	894	945	1 047	994	930	932	8 734
	Comm° congrès publiée	166	178	129	173	152	233	145	179	195	215	1 765
<b>Congrès sans publication</b>	Communication	571	594	704	629	589	780	770	676	561	534	6 408
	Poster	6	6	13	15	29	32	49	50	88	60	348
<b>Chap. ouvrage</b>	Chapitre d'ouvrage	472	474	592	580	720	760	729	762	578	575	6 242
<b>Dir. ouvrage</b>	Direction d'ouvrage, Proceedings, Dossier	107	95	115	123	120	136	116	101	115	100	1 128
<b>Ouvrage</b>	Ouvrage (yc édition critique et traduction)	80	102	96	91	107	94	120	116	109	86	1 001
<b>Autre publié</b>	Autre publication	123	179	181	155	121	144	126	148	138	124	1 439
	Brevet										1	1
<b>Thèse / HDR</b>	Thèse	44	34	32	38	37	43	75	74	81	75	533
	HDR	2	4	5	4		1	6	4	16	5	47
<b>Données de la recherche</b>	Image		19	2	12	1		7		1	2	44
	Son								4	9	1	14
	Vidéo	8	7	4		4	1		1	1	1	27
<b>Autres, non publié</b>	Cours								2		3	5
	Mémoire d'étudiant	1	2	3	6	5	4	10	77	103	61	272
	Autre rapport, séminaire, workshop...							2				2
	Rapport	29	37	48	59	65	122	122	133	116	124	855
	Pré-publ., doc. de travail	43	40	35	23	29	22	41	34	47	45	359
<b>Total général</b>	<b>2 260</b>	<b>2 535</b>	<b>2 699</b>	<b>2 788</b>	<b>2 873</b>	<b>3 317</b>	<b>3 365</b>	<b>3 355</b>	<b>3 088</b>	<b>2 944</b>	<b>29 224</b>	
	<b>Dont publications</b>	<b>1 556</b>	<b>1 792</b>	<b>1 853</b>	<b>2 002</b>	<b>2 114</b>	<b>2 312</b>	<b>2 283</b>	<b>2 300</b>	<b>2 033</b>	<b>20 310</b>	

Source : HAL, traitements OST

Le graphique 9 indique que les productions publiées représentent environ 70 % du corpus MC-HAL. Trois types de productions dominent dans le corpus: les articles (36 %), les communications dans un congrès non publiées (23 %) et les chapitres d'ouvrage (21 %). Si tous les types de participation à un ouvrage sont agrégées, les ouvrages constitueraient le deuxième type avec 28 %.

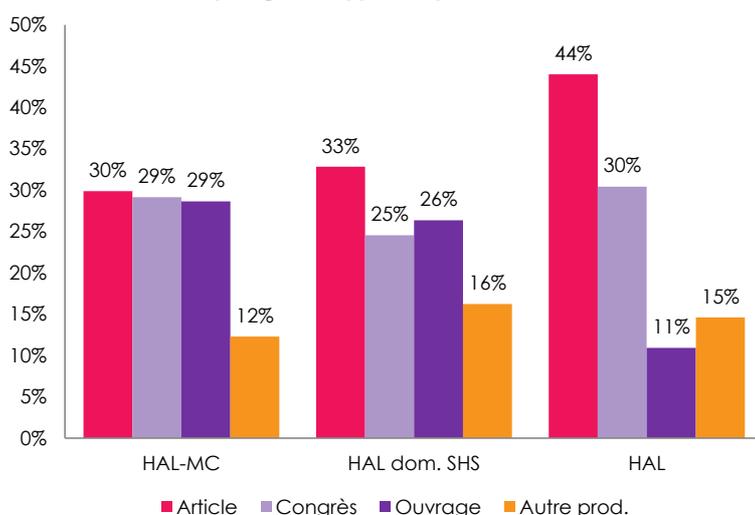
**Graphique 9 : Répartition des types de production – HAL-MC**



Source : HAL, traitements OST

Le regroupement en 4 ensembles (graphique 10), articles, congrès, ouvrages et autres productions (voir Annexe II) permet de comparer la composition du corpus avec les dépôts dans HAL plus généralement. Les actes de congrès ne sont pas une catégorie disponible dans HAL : cette comparaison agrège dans un type Congrès ceux qui ont donné lieu à publication d'actes et les autres. Le graphique 10 montre ainsi que le profil des productions du corpus MC-HAL est proche de celui des dépôts SHS dans HAL. Les articles ont une part proche de 30 %, les ouvrages et les communications dans des congrès (avec actes ou pas) une part comprise entre 25 et 30 %. La part des autres productions est plus élevée dans le corpus HAL dom. SHS que dans le corpus HAL-MC. Si les unités concernées ont la même propension à déposer et que les deux corpus sont également représentatifs (ce qui ne peut être vérifié), cela signifierait que les unités MC ont en fait une part d'autres productions inférieure aux disciplines SHS en général. Il faudrait analyser par exemple les rapports de fouilles ou les catalogues d'exposition.

**Graphique 10 : Distribution par grand type de production dans MC-HAL, HAL dom. SHS et HAL**



Source : HAL, traitements OST

L'ensemble des dépôts dans HAL comporte une part d'articles de 50% supérieur à 44 %. La part des congrès est en revanche similaire et c'est la part des ouvrages qui est sensiblement inférieure à 11 %. Cette différence est sans doute due aux disciplines hors SHS où les ouvrages ont une importance moindre dans la production scientifique et où les ouvrages de vulgarisation sont sans doute aussi moins développés.

### 3 / Langue de production

La langue de production est un élément important de caractérisation : l'utilisation de la langue française est un marqueur des SHS françaises.

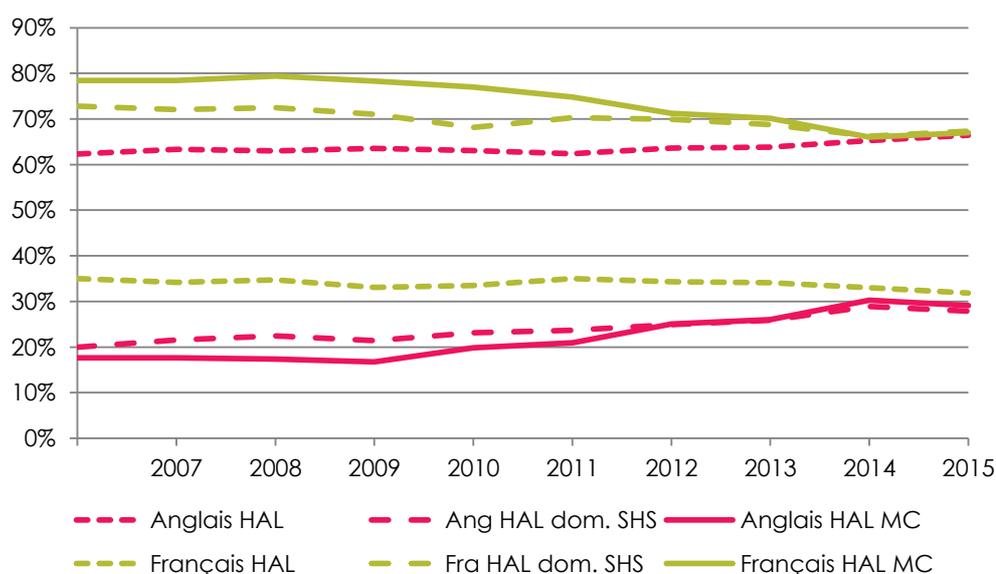
**Tableau 10 : Nombre de productions par langue dans le corpus HAL-MC, 2007-16**

Total 2007-16		Total 2007-16		Total 2007-16	
Français	21 540	Turc	14	Arabe	1
Anglais	6 554	Catalan	11	Bengali	1
Italien	446	Russe	9	Grec moderne	1
Espagnol	241	Tchèque	7	Occidentale	1
Allemand	221	Japonais	7	Coréen	1
Portugais	71	Croate	6	Lituanien	1
Indéfini	44	Latin	5	Romanche	1
Bulgare	15	Polonais	4	Slovène	1
Roumain	15	Néerlandais	3	Albanais	1
		Afar	1	Suédois	1

Source : HAL, traitements OST

L'utilisation du français baisse progressivement sur l'ensemble de la période mais s'estompe sur la fin de la période. La progression de l'anglais est inverse.

**Graphique 11 : Évolution des langues utilisées dans les productions MC-HAL // HAL dom. SHS // HAL**

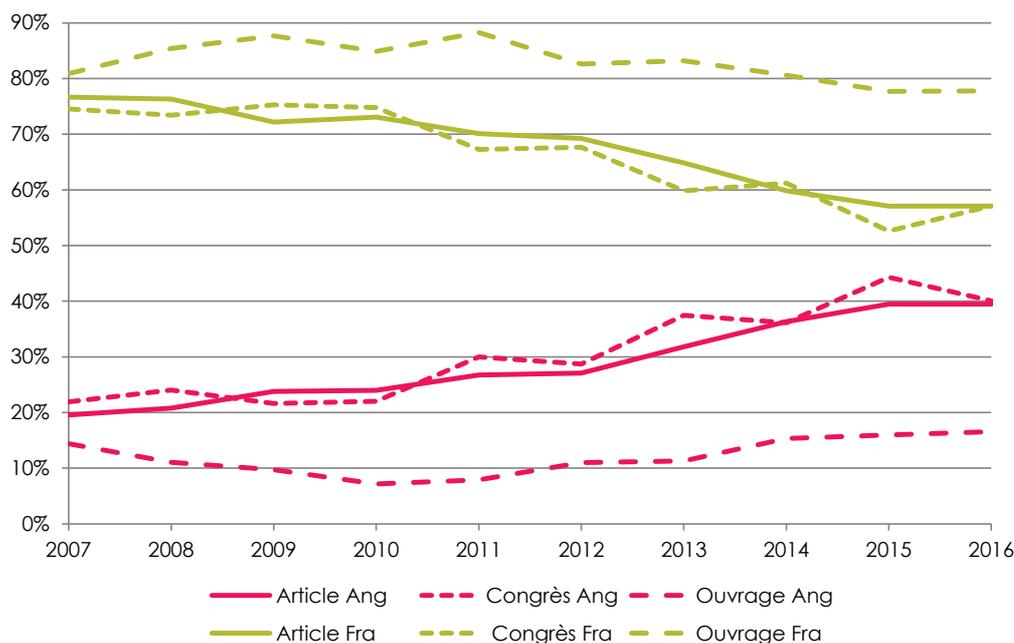


Source : HAL, traitements OST

Sur les 3 corpus, on différencie aisément une utilisation contrastée des deux langues dans les productions scientifiques. Alors que l'anglais est utilisé à plus de 60 % dans le corpus HAL, c'est le français qui prédomine pour les corpus HAL dom. SHS (70 %) et MC-HAL (80 %). En fin de période cependant la part du français dans MC-HAL baisse en dessous de 70 %.

Les pratiques varient selon les supports de production. Les communications dans un congrès, ainsi que les articles, sont publiés en anglais à 30 % sur l'ensemble de la période. Les autres productions, dont les ouvrages, sont en français à 90 % ou plus. L'évolution sur la période a aussi été très différenciée par type de production. Les articles et actes de colloque sont désormais publiés à 40 % en anglais, alors que les ouvrages sont en anglais dans moins de 20 % des cas.

**Graphique 12 : Part du français et de l'anglais dans le corpus MC-HAL, par type de production**



Source : HAL, traitements OST

## 4 / Domaines de recherche et thématiques

### Distribution du corpus par domaine de recherche

HAL repose sur la participation des déposants. Ces derniers renseignent les différents champs d'information et, notamment, le ou les domaines scientifiques de rattachement de la production déposée.

La classification mise à disposition comprend 13 grands domaines :

- Chimie,
- Économie et finance quantitative,
- Informatique,
- Mathématiques,
- Physique,
- Planète et Univers,
- Science non linéaire (physique),
- Sciences cognitives,
- Sciences de l'environnement,
- Sciences de l'Homme et Société,
- Sciences de l'ingénieur,
- Sciences du Vivant,
- Statistiques.

Sur les 29 224 productions HAL-MC, 649 n'ont pas d'indexation de domaine. Sur les 28 575 productions ayant une indexation de domaine, 21 413 n'ont qu'un seul domaine de rattachement (mono-disciplinaire) ; soit plus de 75 % du corpus, alors que sur HAL-SHS, ce taux n'est que de 69 %. Chaque grand domaine comporte des domaines que les déposants peuvent sélectionner sans limite de nombre de domaines. Cette indexation par l'utilisateur est variée et variable.

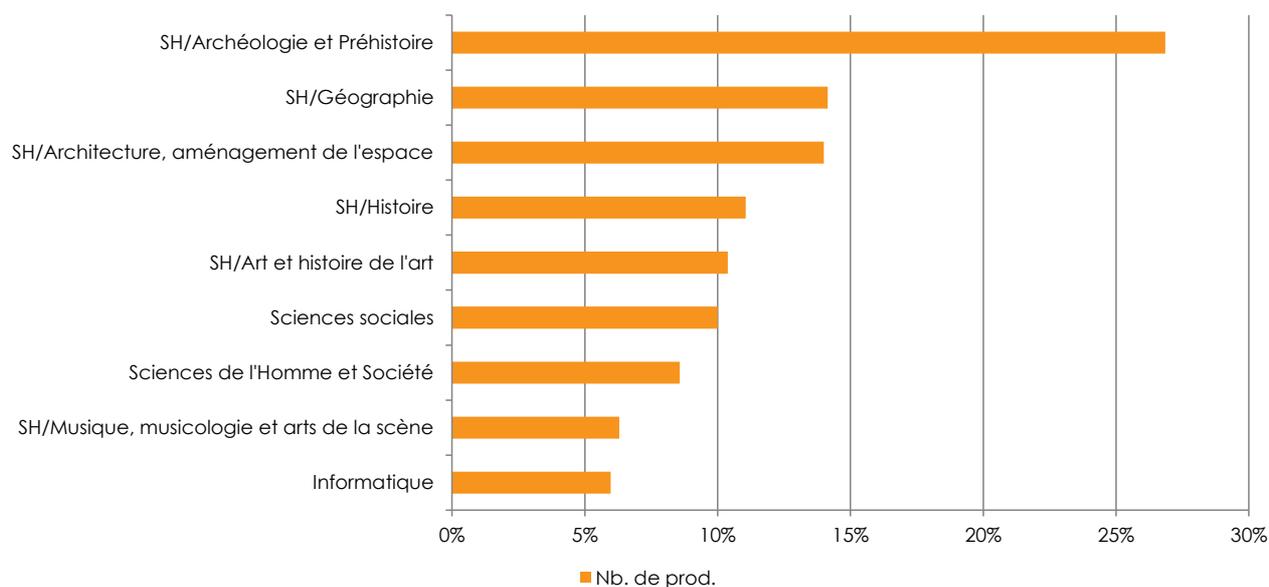
Exemples :

Domaine :

<b>Domaine :</b>	<b>Physique [physics] / Physique [physics] / Optique [physics.optics]</b> Physique [physics] / Physique [physics] / Chimie-Physique [physics.chem-ph] Sciences de l'Homme et Société / Héritage culturel et muséologie
<b>Domaine :</b>	<b>Sciences de l'Homme et Société / Art et histoire de l'art</b> Sciences de l'Homme et Société / Anthropologie sociale et ethnologie Sciences de l'Homme et Société / Histoire, Philosophie et Sociologie des sciences Sciences de l'Homme et Société / Héritage culturel et muséologie Sciences de l'Homme et Société / Sciences de l'information et de la communication Sciences de l'Homme et Société / Histoire

Au sein du corpus MC-HAL, les Sciences de l'homme et Société (SHS) sont mentionnées pour 25 604 items déposés, soit 90 % des productions indexées. De nombreux dépôts contiennent une dimension liée aux SHS même s'ils sont issus d'une autre discipline. Afin d'avoir une vision plus claire du corpus, pour les productions traitant des domaines autres que les SHS, les informations ont été agrégées. Ainsi, pour une production traitant de Chimie/Chimie analytique et Chimie/Autre, seul le niveau 'Chimie' a été retenu. Pour les productions relevant des SHS, le niveau le plus fin a été conservé en dissociant les Sciences sociales des Sciences humaines : pour une production traitant de Sciences humaines/Archéologie et Préhistoire, ce niveau d'information a bien été conservé. Le corpus a donc été analysé sur 32 domaines au lieu de 297, facilitant ainsi la lecture des données tout en conservant des précisions concernant les sciences humaines. Malgré ces regroupements, la multiplicité des indexations possibles dans HAL ne permet pas de détailler l'évolution sur la période et la distribution par domaines disciplinaire est analysée pour l'ensemble de la période.

**Graphique 13 : Les 9 premiers domaines disciplinaires du corpus MC-HAL, 2007-16**



Source : HAL, traitements OST

Les 9 premiers domaines précisés sur le graphique 13 représentent 24 971 productions, soit 85 % du corpus MC-HAL. Le domaine archéologie/préhistoire représente 27 % du corpus. Les domaines de la géographie et l'architecture/aménagement de l'espace en représentent 14 % et l'histoire, les sciences sociales et l'art / l'histoire de l'art respectivement 11 %. Cette distribution est difficilement comparable avec celle du corpus MC-WoS dans la mesure où la classification des domaines est différente, avec des intitulés et sans doute des périmètres qui diffèrent. Ainsi, dans HAL, l'architecture est-elle groupée avec l'aménagement de l'espace, ce qui n'est pas le cas dans la nomenclature du WoS.

Les productions indexées en sciences de l'homme et de la société sans autre précision, en musique, musicologie et arts de la scène à égalité et en sciences de l'ingénieur représentent chacune moins de 10 % du corpus.

**Tableau 11 : Indices de spécialisation au sein de HAL, par domaine ayant plus de 200 productions entre 2007 et 2016**

	2007-2011	2012-2016
Informatique/Son	21,0	31,2
<b>Sc. humaines</b> /Archéologie et Préhistoire	17,2	28,6
<b>Sc. humaines</b> /Architecture, aménagement de l'espace	10,5	27,9
<b>Sc. humaines</b> /Musique, musicologie et arts de la scène	10,7	20,2
<b>Sc. humaines</b> /Art et histoire de l'art	10,6	15,6
<b>Sc. humaines</b> /Géographie	5,1	12,0
<b>Sc. humaines</b> /Études de l'environnement	8,6	11,8
Planète et Univers/Sciences de la Terre/Géomorphologie	4,7	11,0
<b>Sc. humaines</b> /Religions	5,7	10,7
<b>Sc. humaines</b> /Études classiques	4,9	9,5
<b>Sc. humaines</b> /Anthropologie biologique	6,2	9,1
Informatique/Multimédia	5,5	7,8
<b>Sc. humaines</b> /Histoire	3,1	6,1
<b>Sc. humaines</b> /Héritage culturel et muséologie	2,8	6,0
Sciences de l'Homme et Société	4,0	5,4
Sc. de l'ingénieur /Acoustique	3,8	5,1
Informatique/Synthèse d'image et réalité virtuelle	2,6	4,7
Informatique/Modélisation et simulation	1,0	4,2
<b>Sc. humaines</b> /Anthropologie sociale et ethnologie	2,9	4,1
Sciences cognitives/Neurosciences	4,0	3,9
Sc. de l'environnement/Environnement et Société	1,9	3,8
<b>Sc. humaines</b> /Sociologie	0,6	3,6
Informatique/Interface homme-machine	3,0	3,6
<b>Sc. humaines</b> /Histoire, Philosophie et Sociologie des sciences	2,1	3,3
<b>Sc. humaines</b> /Littératures	1,7	2,9
<b>Sc. humaines</b> /Philosophie	1,7	2,6
<b>Sc. humaines</b> /Sciences de l'information et de la communication	0,6	1,8
Sc. de l'ingénieur /Traitement du signal et de l'image	1,2	1,8
Sciences de l'environnement	0,8	1,8
Sc. de l'environnement/Milieus et Changements globaux	1,0	1,8
Informatique/Traitement du signal et de l'image	0,9	1,4
Informatique/Intelligence artificielle	1,5	1,1
<b>Sc. humaines</b> /Linguistique	0,4	1,1
Informatique	0,5	0,6

Source : HAL, traitements OST

Les indices supérieurs à 1 indiquent les domaines dans lesquels le MC a une activité plus forte que l'ensemble des productions déposées dans HAL.

Les indices des 5 domaines ayant les indices les plus élevés dépassent 15 en seconde période. Le premier domaine appartient à l'informatique (son) et les quatre suivants aux sciences humaines : archéologie et préhistoire, architecture et aménagement de l'espace, musique/musicologie/ arts de la scène, art/ histoire de l'art, géographie.

Ces indices très élevés pourraient avoir deux types d'explication. Premièrement, les structures MC pourraient être les seules à produire dans le domaine, comme par exemple l'informatique/son avec l'UMR à laquelle participe l'IRCAM. Deuxièmement, les dépôts dans HAL étant volontaires, l'archive pourrait ne présenter qu'une petite partie de la production française dans certains domaines. Notons d'ailleurs que quatre autres domaines ont des indices à plus de 10, y compris sciences de la terre/géomorphologie.

Parmi les domaines les plus spécialisés du corpus MC-HAL, quatre étaient des domaines de spécialisation déjà identifiés dans le corpus MC-WoS : architecture, archéologie, géographie et arts. Les périmètres précis peuvent différer, mais des domaines communs sont bien identifiés (graphique 4).

## Cartographie des thèmes de recherche dans le corpus MC-HAL

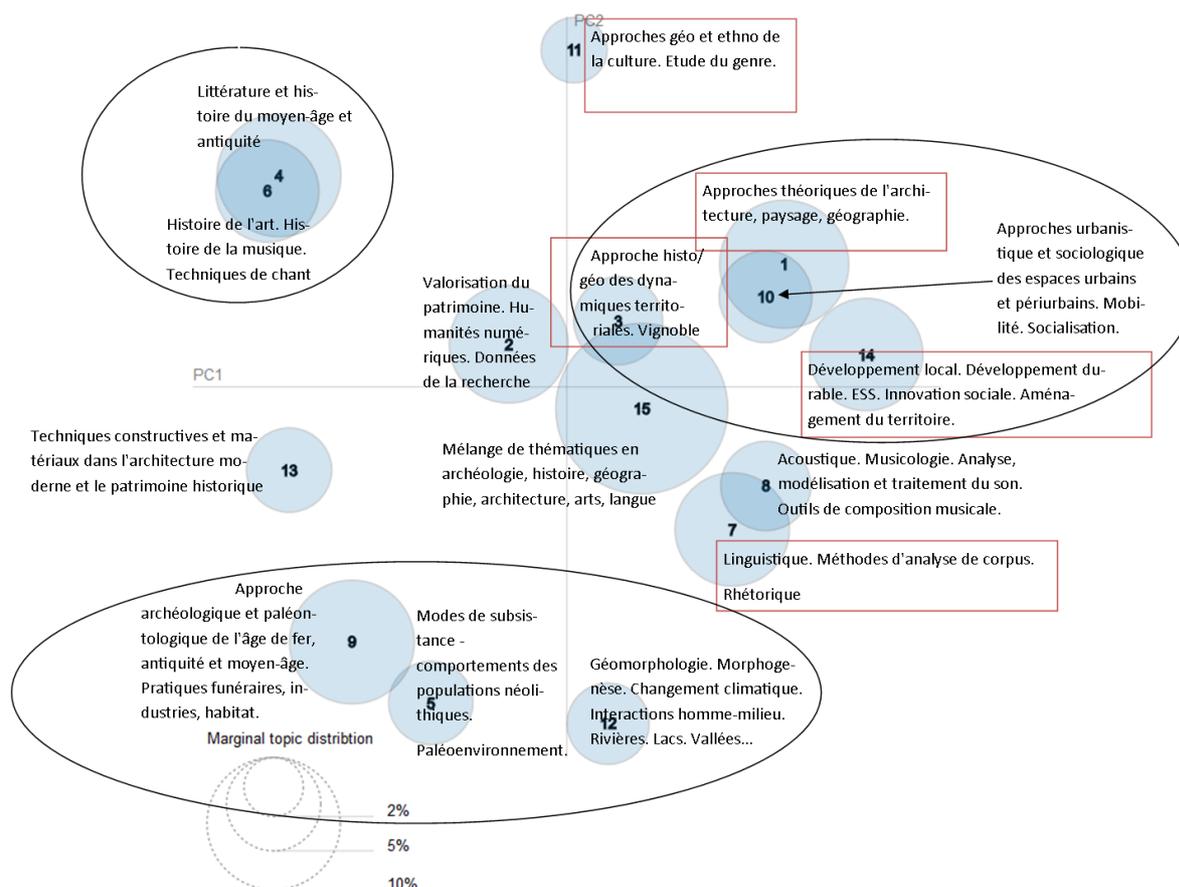
L'analyse thématique a été conduite avec la même méthode que sur le corpus MC-WoS (voir p. 9). Ici, seuls les dépôts disposant d'un résumé en français ont été mobilisés, soit 7 757 productions sur 29 224, soit 27 % du corpus seulement. Les dépôts ayant un résumé dans une autre langue ont été ignorés.

La cartographie dessine des clusters (ensembles de bulles liées entre elles) qui ont ensuite été nommés à partir d'une analyse des mots les plus fréquents, et de la lecture des titres et résumés correspondants. Cette analyse n'a pas fait appel à des experts du domaine et il ne s'agit donc que d'une première approche qui devrait être validée et complétée.

Par rapport à la cartographie issue du WoS, les thématiques apparaissent plus dispersées et la cartographie est plus délicate à interpréter. La thématique 15 regroupe un « mélange » de thèmes. Malgré une composition en termes de documents différente, on retrouve cependant les grands thèmes identifiés sur le précédent corpus.

Dans HAL, les objets archéologiques (bulles 5, 9,12) paléontologiques et technologiques apparaissent moins fortement ou moins nettement que ces thèmes dans le corpus MC-WoS. Le corpus MC-HAL porte principalement sur des objets d'études en urbanisme (1, 10), en architecture et en géographie (1, 15). Les thématiques de développement local (14) et d'aménagement du territoire (3) apparaissent, alors qu'elles étaient absentes dans le corpus WoS-MC. Les thématiques de l'histoire de l'art (6) de l'histoire de la musique (4) et d'acoustique (8) sont présentes comme dans le WoS. La culture et le genre (11) comme objet d'étude font leur apparition, ainsi que la linguistique (7).

Figure 4 : Cartographie des thématiques du corpus MC - HAL par Topic Modeling



Source : HAL, traitements OST

Remarque : les nombres ne servent qu'à repérer les différents thèmes, ils n'ont pas d'autre signification

## ANALYSE DU CORPUS MC-HCÉRES

Le corpus MC-Hcéres étant constitué des unités évaluées par le Hcéres, il ne concerne que ces dernières ; les productions des unités non évaluées n'y figurent pas. Donc, a priori, les productions sont exhaustives pour les unités évaluées et non existantes pour les unités non évaluées. Ces dernières comportant les UMR, elles représentent cependant une majorité de la production.

Les informations fournies par chaque structure ne concernent que les années précédant cette évaluation et ne couvrent que 5 années glissantes en fonction des vagues d'évaluation, soit environ la moitié de la décennie couverte par l'analyse pour les deux autres corpus.

### 1 / Nombre et types de production

Il n'existe pas de base de données reprenant les informations contenues dans les rapports d'autoévaluation des structures évaluées par le Hcéres. Un travail spécifique a été réalisé afin d'extraire les informations des dossiers et de les mettre dans un format exploitable à des fins statistiques.

Compte tenu des dates indiquées dans les dossiers d'autoévaluation, l'ensemble des informations est pris en compte, les productions ayant des dates non comprises dans la période d'analyse de 2007 à 2016 représentant moins de 4 % de la production (tableau 12).

**Tableau 12 : Période de couverture des données MC-Hcéres**

	Nombre de productions	Parts
<b>Avant 2007</b>	379	0,9%
<b>2007-2016</b>	41 616	96,3%
<b>2017-2018</b>	10	0,0%
<b>À paraître, soumis, sous presse, sans date...</b>	1 195	2,8%
<b>Nombre total de productions</b>	<b>43 200</b>	<b>100%</b>

Source : Hcéres, traitements OST

La liste des types de production fournie par le Hcéres dans le cadre des évaluations est large. Elle a évolué au avec les différentes vagues. Le tableau 13 précise les éléments communs permettant de caractériser la production de la recherche culturelle du MC.

**Tableau 13 : Libellés et codes des productions scientifiques suivant la nomenclature du Hcéres**

Produit	Code	Description
<b>Article</b>	ACL	Articles dans des revues internationales ou nationales avec comité de lecture répertoriées par le Hcéres ou dans les bases de données internationales (ISI Web of Knowledge, Scopus, Pub Med...)
	ACLN	Articles dans des revues avec comité de lecture non répertoriées par le Hcéres ou dans des bases de données internationales
	ASCL	Articles dans des revues sans comité de lecture
<b>Actes</b>	ACTI	Communications avec actes dans un congrès international
	ACTN	Communications avec actes dans un congrès national
<b>Communication sans publication</b>	INV	Conférences données à l'invitation du Comité d'organisation dans un congrès national ou international
	COM	Communications orales sans actes dans un congrès international ou national
	AFF	Communications par affiche dans un congrès international ou national
<b>Ouvrage</b>	OS	Ouvrages scientifiques
	OV	Ouvrages de vulgarisation
	DO	Directions d'ouvrages ou de revues
	OS-chap	Chapitres dans ouvrages scientifiques
	OV-chap	Chapitres dans ouvrages de vulgarisation
<b>Autre produit</b>	AP	Bases de données, logiciels enregistrés, traductions, comptes rendus d'ouvrages, rapports de fouilles, guides techniques, catalogues d'exposition, rapports intermédiaires de projets internationaux...

Source : Hcéres, traitements OST

La nomenclature de productions du Hcéres est plus détaillée que des bases de données internationales ou même que celle proposée dans HAL. Elle distingue notamment les articles publiés dans les revues sans comité de lecture et les ouvrages de vulgarisation. Cela correspond à l'objectif de l'évaluation de prendre en compte l'ensemble des activités et production des unités de recherche évaluées. Lors de la rédaction des rapports d'autoévaluation, la qualification d'une production est déclarative et repose donc sur la compréhension des catégories de la nomenclature par les évalués. La catégorie autre produit rassemble des documents variés, comprenant à la fois des produits de la recherche et des rapports intermédiaires.

Dans la nomenclature Hcéres, les actes de colloques, internationaux (ACTI) ou nationaux (ACTN) publiés sont distincts des simples communications dans un congrès. Ils ont donc été regroupés avec les articles afin d'obtenir une catégorie unique, comparable à celle utilisée pour les deux autres corpus. De même, afin de rapprocher les différentes nomenclatures, les ouvrages de vulgarisation ont été regroupés avec les autres produits.

**Tableau 14 : Répartition des types de production MC-Hcéres**

	Type de production*	Nombre	Part du type de production
<b>Articles</b>	Article de revue	<b>8 500</b>	<b>22,0%</b>
	ACL	5 874	13,6%
	ASCL	2 113	4,9%
	ACLN	1 513	3,5%
	Congrès avec actes	<b>6 107</b>	<b>14,2%</b>
	ACTI	4 483	10,4%
	ACTN	1 624	3,8%
<b>Congrès</b>	Congrès sans actes	<b>8 597</b>	<b>19,9%</b>
	COM	6 402	14,8%
	INV	1 589	3,7%
	AFF	606	1,4%
<b>Ouvrages</b>	Ouvrage scientifique	<b>7 954</b>	<b>18,4%</b>
	OS-Chap	3 716	8,6%
	OS	2 993	6,9%
	DO	1 245	2,9%
<b>Autre production</b>	Autre production	<b>11 042</b>	<b>25,6%</b>
	AP	9 833	22,8%
	OV	739	1,7%
	OV-Chap	470	1,1%
	<b>Total</b>	<b>43 200</b>	<b>100%</b>

Source : Hcéres, traitements OST

\* Voir libellés tableau 13

Les articles publiés dans des revues avec comité de lecture répertoriées par le Hcéres ou dans les bases de données internationales forment la catégorie la plus importante (ACL) avec près de 14 % du corpus. Les actes de congrès internationaux (ACTI, 10,4 %) et les actes de congrès nationaux (ACTN, 3,8 %), publiés, peuvent être rapprochés des articles et complètent cette catégorie qui se rapproche du périmètre des articles et actes des deux autres bases. Cet ensemble représente 28 % du corpus Hcéres. Les autres articles, publiés dans des revues sans comité de lecture (ASCL, 5 %) ou dans des revues avec comité de lecture mais non répertoriées dans des bases de données internationales (ACLN, 3,5 %), représentent un peu moins de 10 % de la production. Les documents non publiés (Congrès, 20 % et Autre produit, 23 %) représentent 43 % du corpus MC-Hcéres.

L'ensemble des ouvrages a été réparti, puisque la nomenclature du Hcéres le permet, en répertoriant d'une part les ouvrages scientifiques et d'autre part les ouvrages de vulgarisation. L'ensemble de ces derniers et des autres produits, essentiels dans certaines équipes, représentent 27 % des productions dans MC-Hcéres. La comparaison directe n'est pas possible avec le corpus MC-HAL dans la mesure où il ne distingue pas les ouvrages de vulgarisation. Il est simplement possible de comparer la part du total des ouvrages dans MC-HAL (28 %) avec la part des deux types d'ouvrages dans MC-Hcéres (21 %). Il faut cependant aussi tenir compte du fait que dans MC-Hcéres, les autres produits ont une part importante par rapport aux autres corpus.

## 2 / Production par type de structure

Les structures évaluées par le Hcéres peuvent être rassemblées en 3 groupes : les EPN (établissement public national), les UMR (unités mixtes de recherche) et les UR (unités de recherche) des écoles nationales d'architecture. Cette nomenclature générique permet de dissocier les productions des UMR de celles des autres structures qui n'ont pas les mêmes missions en matière de production scientifique.

**Tableau 15 : Répartition de la production par type et par structure**

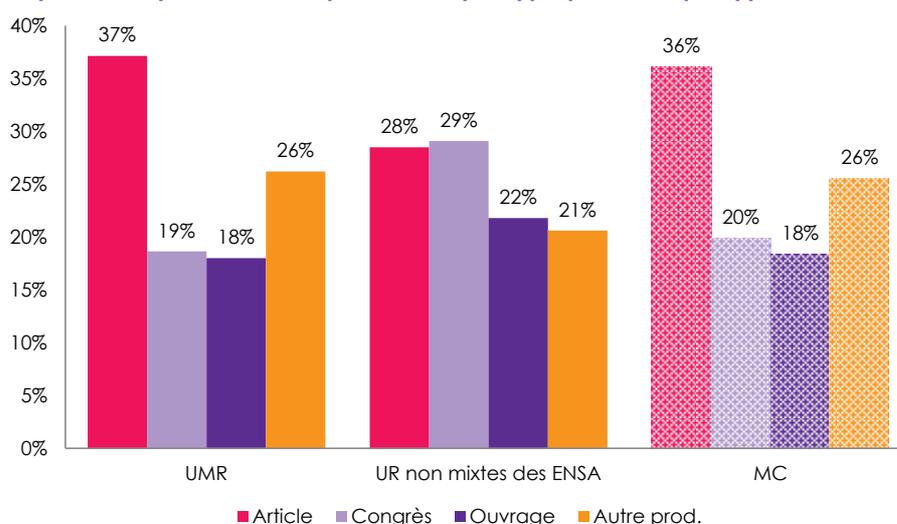
	Article	Actes/Congrès	Ouvrage	Autre prod.	MC global	Part
<b>EPN</b>		54			54	0,1%
<b>UMR</b>	14 239	7 146	6 907	10 052	38 344	88,8%
<b>UR non mixtes des ENSA</b>	1 368	1 397	1 047	990	4 802	11,1%
<b>MC global</b>	<b>15 607</b>	<b>8 597</b>	<b>7 954</b>	<b>11 042</b>	<b>43 200</b>	

Source : Hcéres, traitements OST

Les EPN ayant déposé un dossier d'autoévaluation n'indiquent que des articles dans leur production scientifique, soit 0,1 % du corpus MC. Les UMR génèrent la presque totalité des productions scientifiques, ce qui est cohérent avec leur vocation de recherche. Les UR non mixtes des ENSA participent à plus de 11 % du corpus MC sur la période considérée.

Le graphique 14 fournit la répartition des productions par type de structure. Les UMR ont une part sensiblement plus élevée d'articles dans leur production (37 %) que les UR. Les participations aux congrès (19 %) sont à l'inverse moins représentées dans la production totale des UMR.

**Graphique 14 : Répartition de la production par type pour chaque type de structure,**



Source : Hcéres, traitements OST

Une partie importante des productions des UMR est regroupée au sein des autres productions (26 %). Ce sous-corpus qui est très hétérogène puisqu'il est composé de bases de données, logiciels enregistrés, traductions, comptes rendus d'ouvrages, rapports de fouilles, guides techniques, catalogues d'exposition, rapports intermédiaires de grands projets internationaux. L'analyse détaillée de cette production apporterait un éclairage complémentaire sur la production de la recherche culturelle.

## 3 / Structures et productions par domaine disciplinaire

Les listes de productions issues des dossiers du Hcéres ne disposent pas de notices comme dans la base WoS ou dans HAL. En conséquence, les métadonnées utilisables pour analyser les thématiques sont moins riches. Il n'y a pas de résumé ou de mots clé en particulier. L'analyse des thèmes traités ne peut donc reposer que sur la seule classification disponible soit celle liée aux structures évaluées. Très générale, elle est donc attribuée à l'ensemble de la structure.

Une première approche a consisté à regrouper les structures par domaine disciplinaire du Hcéres. Cette distribution a ensuite été comparée à la distribution des productions par domaine disciplinaire. Les tableaux 16 et 17 permettent de comparer les deux distributions.

Tableau 16 : Répartition des structures évaluées par domaine disciplinaire Hcéres

Domaine disciplinaire Hcéres	Part du total évalué	Intitulé de l'unité
<b>SHS 3</b> Espace, environnement et sociétés	53%	UMR 1563 AAU Ambiances, Architectures, Urbanités
		UMR 3329 AUSser Architecture, Urbanisme, Société : savoirs, enseignements, recherche
		UMR 3495 MAP Modèles et simulations pour l'architecture et le patrimoine
		UMR 5319 PASSAGES
		UMR 5600 EVS Environnement, Ville, Société
		UMR 7218 LAVUE Laboratoire Architecture, Ville, Urbanisme, Environnement
		AE&CC Architecture, Environnement et Cultures Constructives
		AMUP Architecture, Morphologie/Morphogenèse Urbaine et Projet
		Architecture, Paysage, Montagne
		ATE Normandie
		EVCAU - Espace Virtuel de Conception Architecturale et Urbaine
		Groupe de Recherche sur l'Invention et l'Évolution des Formes
		GSA Laboratoire géométrie structure architecture
		INAMA - Investigations sur l'histoire et l'actualité des mutations architecturales
		LCTH Laboratoire Conception Territoire Histoire
		LéaV - Laboratoire de recherche de l'École nationale supérieure d'architecture de Versailles
		LHAC Laboratoire d'histoire d'architecture contemporaine
		LIAT Laboratoire Infrastructures, Architecture, Territoire
		LIFAM Laboratoire Innovation Formes Architectures Milieu
		LRA Laboratoire de recherche en architecture
MHAevt Métiers de l'histoire de l'architecture, édifices-villes-territoires		
PAVE Profession Architecture Ville Environnement		
TRANSFORMATIONS		
UR Projects		
<b>SHS 6</b> Mondes anciens et contemporains	26,5%	Musée du Quai Branly - Dept Recherche
		UMR 5140 ASM Archéologie des sociétés méditerranéennes
		UMR 5199 PACEA De la Préhistoire à l'Actuel : culture, environnement et anthropologie
		UMR 5608 TRACES Travaux et recherches archéol. sur les cultures, les espaces et les sociétés
		UMR 6298 ArTeHis Archéologie, Terre, Histoire, Sociétés
		UMR 6566 CREAHA Centre de recherche en archéologie, archéosciences, histoire
		UMR 7041 ArScAn Archéologies et sciences de l'antiquité
		UMR 7044 ARCHIMEDE Archéologie et histoire ancienne : Méditerranée et Europe
		UMR 7269 LAMPEA Laboratoire méditerranéen de Préhistoire Europe-Afrique
		UMR 7302 CESC Centre d'études supérieures de civilisation médiévale
		UMR 8150 Centre André Chastel
		UMR 8164 HALMA Histoire, archéologie, littérature des mondes anciens
		USR 3103 IN VISU - L'info. visuelle et textuelle en histoire de l'art : nvx terrains, corpus, outils
<b>Sciences et technologies</b> Maths, Physique, Sc. de la terre et de l'univers, Chimie, Sc. pour l'ingénieur, S&T de l'info et de la comm°	8,2%	UMR 7190 Institut Jean Lerond D'Alembert - LAM Lutheries - Acoustique - Musique
		UMR 9912 STMS Sciences et technologies de la musique et du son
		USR 3224 CRC Centre de recherche sur la conservation
<b>SHS 5</b> Langues, textes, arts et cultures	8,2%	USR 3461 IPANEMA Institut photonique d'analyse non destructive européenne des matériaux anciens
		partenaire de Scènes du monde, création, savoirs critiques
		UMR 7323 CESR Centre d'études supérieures de la Renaissance
		UMR 8218 ACTE Arts, Créations, Théories, Esthétiques
<b>SHS 2</b> Normes, institutions & comportements sociaux	2%	UMR 8223 IReMus Institut de recherche en musicologie
		UMR 8177 - IIAC Institut Interdisciplinaire d'Anthropologie du Contemporain - LAHIC Laboratoire d'Anthropologie et d'Histoire de l'Institution de la Culture
<b>SHS 4</b> Esprit humain, langage, éducation	2%	UMR 7270 LLL Laboratoire ligérien de linguistique

Source : Hcéres, traitements OST

Certains domaines disciplinaires sont peu représentés : SHS2 et SHS4 n'ont qu'une structure évaluée et représentent 2 % des structures chacun. Le tableau 17 présente les productions par domaine disciplinaire. Les parts de production des différents domaines disciplinaires s'expliquent en partie par la typologie des unités. En effet, les UMR tendent à produire plus que les UR non mixtes des ENSA, tous types de productions confondus.

SHS3 avec 53 % des structures, composées à un tiers d'UMR et deux tiers d'UR des ENSA, participe à 26 % de la production. Symétriquement, SHS6 représente 53 % de la production avec 26,5 % des structures, très majoritairement des UMR.

**Tableau 17 : Répartition des productions scientifiques par domaines disciplinaires du Hcéres, 2007-16**

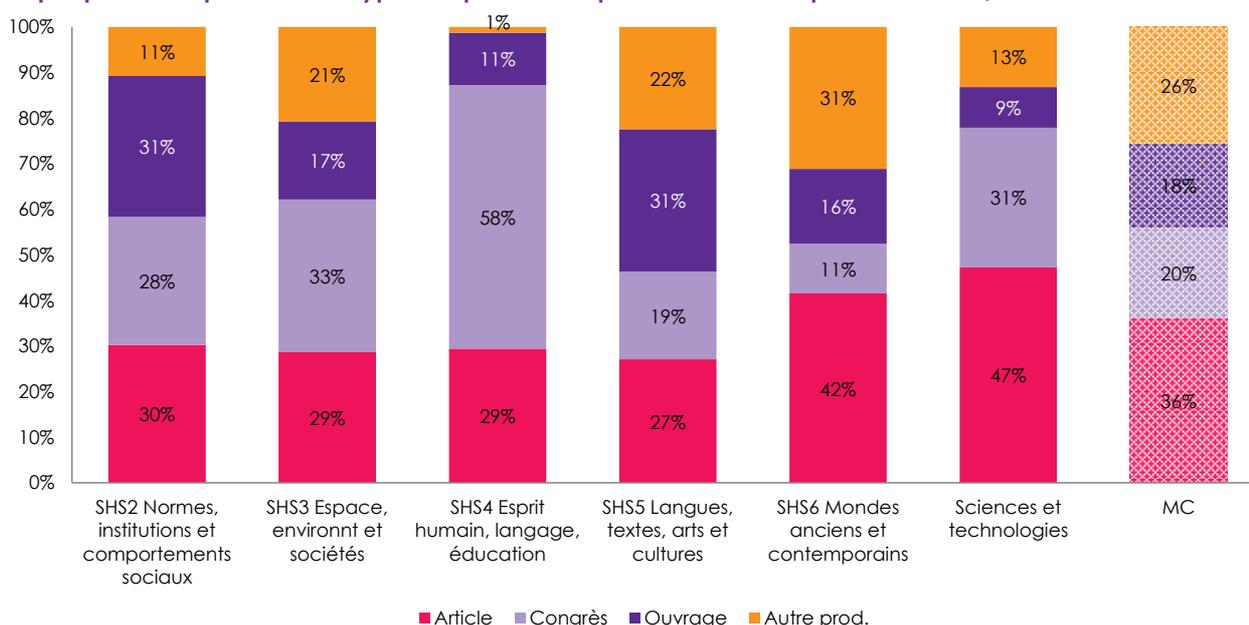
Domaine disciplinaire Hcéres	Sous-domaines disciplinaires	Nb. de prod.	Part
<b>Mondes anciens et contemporains</b>	Histoire, Histoire de l'art, Archéologie	22 803	53%
<b>Espace, environnement et sociétés</b>	Géographie, Aménagement et urbanisme, Architecture	11 066	26%
<b>Langues, textes, arts et cultures</b>	Langues / littératures anciennes et françaises, littérature comparée, Littératures et langues étrangères, Civilisations, Cultures et langues régionales, Arts, Philosophie, sciences des religions, théologie	5 695	13%
<b>Sciences et technologies</b>	Mathématiques, Physique, Sciences de la terre et de l'univers, Chimie, Sciences pour l'ingénieur, S&T de l'information et de la communication	1 951	5%
<b>Normes, institutions et comportements sociaux</b>	Droit, Sc. politique, Anthropologie et ethnologie, Sociologie, Démographie, Sc. de l'information et de la communication	909	2%
<b>Esprit humain, langage, éducation</b>	Linguistique, Psychologie, Sc. de l'éducation, Sc. et techniques des activités physiques et sportives	776	2%
<b>Total MC- Hcéres</b>		<b>43 200</b>	<b>100%</b>

Source : Hcéres, traitements OST

#### 4 / Types de production selon les domaines disciplinaires

Au sein de chaque domaine disciplinaire, les pratiques de production varient sensiblement (graphique 15). La part des articles et actes de congrès publiés est la plus élevée en sciences et technologies, où la part des ouvrages est la plus faible. A l'inverse, la part des articles est la plus faible en langues, textes, arts et cultures où à la fois la part des ouvrages et la part des autres produits sont parmi les plus élevés.

**Graphique 15 : Répartition des types de production par domaine disciplinaire Hcéres, 2007-16**



Source : Hcéres, traitements OST

C'est dans le domaine mondes anciens et contemporains que la part des autres produits est la plus élevée, atteignant près d'un tiers des productions déclarées dans les dossiers Hcéres. Il serait intéressant d'étudier la composition de ces autres produits, par exemple pour évaluer la part qu'y prennent les carnets de fouilles et les catalogues d'exposition.

On voit ainsi apparaître des pratiques de production différenciées selon les domaines disciplinaires. La part des conférences est bien moindre pour les structures relevant de « SHS5 Langues, textes, arts et cultures », à contrario des structures relevant des « SHS4 Esprit humain, langage, éducation ».

## COMPARAISON DES TROIS CORPUS

### 1 / Types de documents et revues selon les corpus

Les analyses des typologies de documents dans les trois corpus permettent une comparaison à partir de grands types. Le tableau 18 souligne que les volumes relatifs des trois corpus dépendent des types de productions pris en compte. Le corpus MC-Hcéres<sup>9</sup> compte 12 fois plus de productions que le corpus MC-WoS. Mais les analyses ont montré que ces corpus sont très hétérogènes, le premier comportant notamment une part importante de productions intermédiaires de la recherche. Si l'on considère des corpus de publications, le ratio n'est plus que de 1 à 6,4, et il descend en dessous de 5 pour les seuls articles et actes de colloques publiés. L'évolution des ratios s'explique par le fait que la part de ces publications dans les corpus MC-HAL et MC-Hcéres est de l'ordre de 36%, contre 93% dans le corpus MC-WoS.

La part des ouvrages est sensiblement plus importante dans MC-HAL que dans MC-Hcéres et il faudrait en analyser précisément la composition, notamment pour distinguer les ouvrages de vulgarisation et pour établir des distinctions par domaine de recherche.

**Tableau 18 : Comparaison des trois corpus MC : productions et part des publications, 2007-16**

	MC-WoS	MC-HAL	MC-Hcéres
Nombre : total de productions	3 594	29 224	43 200
<i>Ratio : nombre / nombre WoS</i>	1	8,1	12,0
Nombre : articles, actes de congrès et ouvrages*	3 346	18 674	23 155
Part dans le corpus	93,1 %	64,6 %	54,6 %
<i>Ratio : nombre / nombre WoS</i>	1	5,3	6,4
Nombre : articles et congrès avec actes publiés	3 346	10 491	15 638
Part : articles et actes de congrès	93,1 %	35,9 %	36,2 %
<i>Ratio : nombre / nombre WoS</i>	1	3,1	4,7

\* Dans HAL, y compris les ouvrages de vulgarisation, les éditions critiques, les traductions et les directions d'ouvrage, proceedings ou dossier.

Dans le corpus Hcéres, hors ouvrages de vulgarisation qui représentent moins de 3 % de ce corpus.

Sources : Clarivate Analytics, HAL & Hcéres, traitements OST

La typologie du Hcéres permet par ailleurs de distinguer les revues avec comité de lecture. Les revues recensées dans le WoS en disposent nécessairement dans la mesure où c'est un critère de sélection. Il serait ainsi intéressant d'approfondir la comparaison en tenant compte de cette distinction pour les trois corpus.

Le tableau 19 compare les revues présentes dans les corpus WoS et HAL pour les unités MC. 414 revues sont présentes dans les deux corpus, ce qui représente 47 % des revues présentes dans MC-WoS et 15 % des revues présentes dans MC-HAL. Symétriquement, 85 % des revues de MC-HAL n'apparaissent que dans ce corpus et 53 % des revues présentes dans MC-WoS ne sont présents que dans ce corpus. Toutes les publications des chercheurs français ayant vocation à être déposées dans HAL, il apparaît une marge de non-dépôt importante (même en tenant compte des contraintes d'embargo).

<sup>9</sup> Comme précisé dans la présentation de la constitution des corpus, le corpus MC-Hcéres porte sur 5 années glissantes, contrairement aux corpus issus du WoS et de HAL qui portent sur 10 ans. L'objectif ici n'est pas de comparer les volumes, mais les volumes relatifs en fonction des types de production pris en compte.

**Tableau 19 : Revues communes aux corpus MC-WoS et MC-HAL**

Revues	Nombre de revues HAL	Nombre de revues WoS	% Revues HAL	% revues WoS
<b>communes aux 2 corpus</b>	414	414	15%	47%
<b>dans HAL seulement, dont</b>	2 381		85%	
<b>dans WoS seulement</b>		461		53%
<b>Total Revues</b>	<b>2 795</b>	<b>875</b>	<b>100%</b>	<b>100%</b>

Sources : Clarivate Analytics & HAL, traitements OST

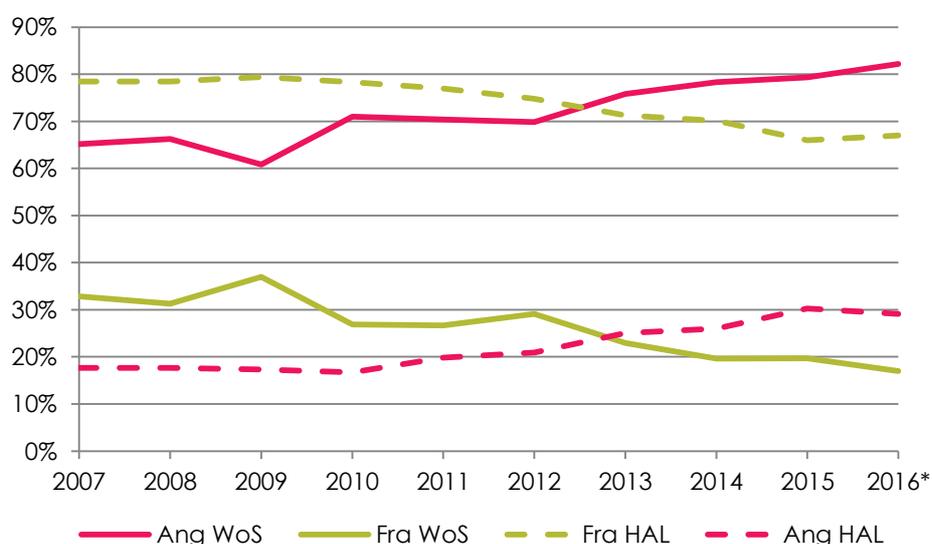
Il serait intéressant d'examiner les revues qui n'apparaissent que dans HAL afin de connaître leur couverture (thématique locale, nationale, internationale), leur mode de publication (comité de lecture ou non) et leur diffusion (papier, numérique) afin d'estimer leur audience. Leur présence ou non dans les listes Hcéres serait aussi un élément d'intérêt pour les caractériser. Par ailleurs, l'identification des revues présentes dans le corpus MC-Hcéres demanderait une étude spécifique.

## 2 / Langues de production

Les parts de l'anglais et du français dans les productions de la recherche culturelle évoluent de manière similaire dans les corpus MC-WoS et MC-HAL. Le graphique 16 souligne que la part du français baisse alors que celle de l'anglais augmente.

La part des productions en anglais atteint 30 % dans le corpus MC-HAL. L'analyse a cependant montré que cette moyenne varie selon les types de productions. Le graphique 12 a montré que la part de l'anglais atteint 40 % pour les articles et les congrès, alors qu'elle est inférieure à 20 % pour les ouvrages. Par ailleurs, il n'existe pas de tendance nette à l'augmentation de cette part dans le cas des ouvrages. Là encore, il serait intéressant de pouvoir distinguer la langue de publication selon les types d'ouvrages en distinguant notamment les ouvrages de vulgarisation.

**Graphique 16 : Évolution des langues dans les publications/productions**



Sources : Clarivate Analytics & HAL, traitements OST

## 3 / Profil disciplinaire des corpus MC-WoS et MC-HAL

Le vocabulaire et les classifications des corpus MC-WoS et MC-HAL ne s'appliquent pas aux mêmes objets (revue pour l'un, différentes productions pour l'autre) et ont des définitions non harmonisées. L'analyse permet toutefois d'indiquer que :

- L'archéologie, l'architecture, l'histoire de l'art et l'histoire, ainsi que la géographie et la paléontologie, sont des thématiques communes aux 2 corpus.
- L'anthropologie et les géosciences ont une présence marquée dans les données du WoS.
- la musique et l'informatique sont plus visibles au sein du corpus HAL.

Les cartographies thématiques fondées sur des corpus différents et surtout sur des vocabulaires différents puisque construit l'une à partir de documents anglais, l'autre de documents en français, présentent des paysages complémentaires :

- Dans le corpus MC-WoS, 4 grands groupes de thèmes ont été identifiés : archéologie/paléontologie, géographie physique, histoire de l'art, technologies et analyses du patrimoine.
- Les thèmes identifiés dans le corpus MC-HAL sont plus éclatés. Outre des thématiques liées à l'archéologie et à la géographie, apparaissent l'urbanisme, l'architecture, la littérature et l'histoire de l'art, ainsi que la musique et la linguistique.

Le corpus MC- Hcéres ne permet pas de conduire ce type d'analyse car l'indexation, liée aux structures, est trop générique.

#### 4 / Indicateurs : spécialisation WoS et HAL ; impact WoS uniquement

Les indices de spécialisation présentent l'activité du MC dans un domaine relativement à une communauté de référence. Il n'a cependant pas la même signification suivant le mode de construction de la base de données : dans le WoS, il donne une information sur l'activité du MC au sein de revues internationales sélectionnées, dans HAL, il indique l'activité du MC au niveau des productions mais dépend des dépôts qui sont volontaires.

Le corpus MC-WoS présente une forte spécialisation en architecture, archéologie, géographie physique, et histoire de l'art avec des valeurs très élevées. Il a une activité plus faible mais constante en paléontologie, chimie analytique et géosciences. Le corpus MC-HAL présente une forte spécialisation en informatique/son, archéologie et préhistoire, architecture/ aménagement de l'espace.

Malgré les différences de construction, il existe donc bien des thématiques communes identifiées sur les deux corpus.

#### 5 / Co-publications et partenariats

Le corpus MC- WoS permet de distinguer les principaux partenaires des structures MC à l'international qui sont les États-Unis et les grands pays européens (Espagne, Italie, Allemagne...). Les réseaux de collaboration avec les autres pays donnent une vision plus complète de ces partenariats : les structures MC ont, comme dans l'ensemble des domaines de recherche, des collaborations plus fréquentes avec les pays d'Europe. Trois grands réseaux de coopération se dessinent à l'échelle mondiale à partir du corpus MC-WoS : l'un centré sur la France, l'autre comprenant l'Italie et le Canada, enfin, celui composé des États-Unis, du Royaume-Uni et de l'Allemagne.

HAL permet en partie ce type d'analyse grâce aux indications d'affiliation des co-publants : la faible qualité des données n'a cependant pas permis leur exploitation.

## CONCLUSIONS ET APPROFONDISSEMENTS

Les trois sources de données exploitées dans le cadre de ce rapport ont permis la création de trois corpus différents en nombre de productions répertoriées, ce qui est en partie dû à des périmètres différents, avec une distinction fondamentale entre productions et publications notamment. Les typologies fournies pour les trois corpus ont en outre souligné la distinction entre les communications à des congrès selon que ceux-ci donnent lieu à publication d'actes ou pas. La distinction est aussi apparue entre revues avec et sans comité de lecture. Enfin, si chaque corpus a un type relatif aux « autres produits », son volume est variable et il apparaît très hétérogène et n'a pas donc pu être analysé en détail.

La base de données bibliométrique WoS est constituée de revues ayant des citations afin de pouvoir développer des indicateurs et des comparaisons internationales, notamment en matière de profil disciplinaire et d'impact scientifique des publications. Ainsi, même si des efforts sont faits pour couvrir l'ensemble des disciplines, les revues de sciences sociales et/ou sciences humaines ne sont prises en compte que dans la mesure où elles reçoivent des citations. L'identification du corpus MC-WoS a permis de positionner les publications par rapport à la France et au monde dans les mêmes domaines de recherche.

L'archive HAL a pour objectif le dépôt de la production nationale par les chercheurs ou leurs institutions. Le dépôt étant volontaire, le recensement de la production est partiel et variable selon les communautés et institutions. Les métadonnées saisies par les déposants permettent de mener des recherches bibliographiques et une approche par domaine disciplinaire. Les données sur les collaborations scientifiques ne sont en revanche pas directement exploitables.

Les listes de productions sont fournies au Hcéres dans le cadre de l'évaluation, et non pas dans le but d'en faire une base de données ou même un fichier de données. Ces listes se veulent exhaustives sur une période de 5 ans, période de l'évaluation, mais leur saisie répond aux contraintes de compréhension et de disponibilité des équipes. Elles recensent notamment différents types de productions qui ne sont pas prises en compte par les bases de données.

L'analyse des trois corpus développés pour caractériser les productions des unités sous tutelle du MC amène à conclure que le choix du corpus le plus approprié dépend de l'objectif visé. Si l'objectif est de répertorier les productions des unités de la manière la plus exhaustive, les dossiers du Hcéres constituent la source la plus complète avec une typologie détaillée. Cependant, toutes les unités ne sont pas évaluées et les données restent insuffisamment structurées. Si l'objectif est de répertorier les publications et de les comparer avec des publications d'institutions étrangères, une base bibliométrique internationale est nécessaire. L'archive HAL permet des comparaisons entre institutions et disciplines mais doit s'accompagner d'une analyse des comportements de dépôt pour essayer d'apprécier la représentativité de ce qui est accessible.

Les informations liées aux disciplines dans les publications/ productions recensées s'appuient sur des vocabulaires différents et ne s'appliquent pas aux mêmes objets. Dans la base WoS, l'indexation fait appel à une nomenclature disciplinaire appliquée au niveau des revues. Dans HAL, le déposant déclare des domaines afin de préciser le contenu du produit déposé. Pour les données issues des dossiers d'évaluation, une nomenclature large des spécialités scientifiques est appliquée par le Hcéres. L'appréhension des contenus à travers des classifications doit en outre tenir compte du fait que selon les corpus, la langue majoritaire est soit le français, soit l'anglais.

### Pistes d'approfondissement

Des informations de caractérisation plus complètes, fiables ou détaillées que celles qui ont été produites dans le cadre de cette étude pourraient être obtenues en améliorant la qualité de chacun des corpus et la richesse des comparaisons.

### Qualité des données des corpus

Les corpus doivent être étudiés avec précision afin de les valider, notamment en vérifiant les affiliations et leurs formes. En effet, la fiabilisation des adresses des unités n'a été prise en charge que par l'OST ; une validation et un enrichissement du fichier des graphies par le MC constitueraient un complément utile pour renforcer la qualité du corpus obtenu.

Il en est de même pour le corpus HAL : une administration efficace d'un portail HAL du Ministère de la culture veillerait à un référencement fiabilisé des structures de recherche et de leurs rattachements institutionnels, une même structure pouvant être déclarée plusieurs fois. Par ailleurs, l'indexation des domaines de recherche est à 75% mono-valuée (1 seul terme) ce qui rend la recherche des productions peu efficace, d'autant plus que les résumés ne sont présents que dans moins de 50% du corpus et que le taux d'archives est très faible (à peine 20% du corpus).

Les productions déclarées au Hcéres n'ont pas fait l'objet d'un processus de validation dans le cadre de cette étude. Les informations déclarées sur les productions, notamment les publications, ont été reprises sans vérification de leur référencement dans un corpus bibliographique de dépouillement de documents publiés (base bibliographique publique ou commerciale). Une fiabilisation des données pourrait donc aussi être étudiée sur ce corpus.

### **Typologies des productions et thématiques**

Une des richesses du corpus MC-Hcéres est la typologie détaillée des productions qu'il propose. L'exploration de ce qui est contenu dans la catégorie Autres productions pourrait affiner la caractérisation des productions des unités concernées en précisant la distribution par type selon les domaines par exemples (bases de données, logiciels enregistrés, traductions, comptes rendus d'ouvrages, rapports de fouilles, guides techniques, catalogues d'exposition...). Les ouvrages de vulgarisation, et plus généralement les productions à destination de la société, pourraient être singularisées et leur impact étudié. La part de ces autres produits est plus faible dans le corpus MC-HAL (12 %, contre 26 % dans MC-Hcéres), ce qui mérite aussi d'être exploré. Ces types d'analyse devront être menés en lien avec les experts des domaines concernés.

Une exploration des revues apparaissant dans les corpus HAL et Hcéres, de leur couverture (thématique locale, nationale, internationale), de leur mode de publication (comité de lecture ou non), leur diffusion (papier, numérique) et, enfin, de leur présence ou non dans les listes Hcéres pourrait être effectuée. Concernant le corpus MC-Hcéres, cela supposerait de mieux structurer les données extraites des dossiers.

Les cartographies thématiques, qui apportent une vision qui n'est plus strictement disciplinaire, mais affichent de nouveaux liens entre domaines de recherche mériteraient un examen approfondi par des experts. Par ailleurs, certaines recherches sont très localisées mais l'indexation thématique omet ce type de caractérisation.

### **Couverture des sources**

La couverture des corpus MC-WoS et MC-HAL pourrait être améliorée. Concernant le WoS, il s'agirait d'examiner la ressource que peut représenter les revues qui sont encore en cours d'examen avant d'avoir été sélectionnées dans la base, mais pour lesquelles des indicateurs sont disponibles (les emerging sources). L'analyse des ouvrages intégrés dans la base pourrait aussi constituer une ressource.

Un effort doit être fait pour assurer des dépôts systématiques dans l'archive ouverte HAL afin de disposer d'une source proche de l'exhaustivité pour connaître, diffuser et caractériser la production des unités MC. Cet effort doit porter non seulement sur l'assistance au dépôt des productions afin que les problèmes techniques soient levés, mais également sur l'administration d'un portail HAL MC, administration technique (support informatique) et documentaire (aide à la saisie et à l'indexation) sans oublier l'axe diffusion et communication. Un tel projet ne pourra se développer sans une volonté et un support fort et affiché de la part du Ministère.

## REFERENCES

- Chi, P. S. (2014a). Which role do non-source items play in the social sciences? A case study in political science in Germany. *Scientometrics*, 101(2), 1195–1213.
- Chi, P. S. (2014b). *The Characteristics and Impact of Non-Source Items in the Social Sciences – A Pilot Study of Two Political Science Departments in Germany*. Berlin: Humboldt-Universität zu Berlin.
- Faggiolani C., Solimine G. (2018) Mapping the Role of the Book in Evaluation at the Individual and Department Level in Italian SSH. A Multisource Analysis. In: Bonaccorsi A. (eds) *The Evaluation of Research in Social Sciences and Humanities*. Springer, Cham
- Hicks, D. (1999). The difficulty of achieving full coverage of international social science literature and the bibliometric consequences. *Scientometrics*, 44(2), 193–215.
- Hicks, D. (2004). The four literatures of social science. In H. F. Moed, W. Glänzel, & U. Schmoch (Eds.), *Handbook of quantitative science and technology research: The use of publication and patent statistics in studies of S&T systems* (pp. 476–496). Dordrecht: Kluwer Academic Publishers.
- Huang, M., & Chang, Y. (2008). Characteristics of research output in social sciences and humanities: From a research evaluation perspective. *Journal of the American Society for Information Science and Technology*, 59(11), 1819–1828.
- Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics*, 106(1), 213–228.
- Nederhof, A. J. (2006). Bibliometric monitoring of research performance in the social sciences and the humanities: A review. *Scientometrics*, 66(1), 81–100.
- Norris, M., & Oppenheim, C. (2007). Comparing alternatives to the Web of Science for coverage of the social sciences' literature. *Journal of Informetrics*, 1(2), 161–169.
- OST, 2018, *La position scientifique de la France dans le monde, 2000-2015*, Hcéres, Paris.
- Sivertsen, G. & Larsen, B. (2012). Comprehensive bibliographic coverage of the social sciences and humanities in a citation index: an empirical analysis of the potential. *Scientometrics*, 91(2), 567–575.
- Waltman, L. (2016). A review of the literature on citation impact indicators. *Journal of Informetrics*, 10(2), 365–391.

## Annexe 1 - Liste des structures prises en compte

Type structure	Tutelle(s)	intitulé unité	Dossier Hcéres
Conseil	MC	CNRA Conseil national de la recherche archéologique	
		Conseil supérieur des archives (France Archives)	
ENSA & ENSA unité de recherche	ENSA Bordeaux (architecture et paysage)	ENSA Bordeaux (architecture et paysage)	
		ARPEGE : Architecture, Recherche et Pratique Expérimentale à Grande Échelle	S2PUR160010346
		CEPAGE Centre de recherche sur l'histoire et la culture du paysage	
		GEVR Groupe d'Études de la Ville Régulière	
		GRECAU : Groupe de Recherche Environnement Conception Architecturale et Urbaine	S2PUR160010346
		PAVE Profession Architecture Ville Environnement	S2PUR160010795
		UMR 5185 ADESS Aménagement, Développement, Environnement, Santé et Sociétés	
	ENSA Bretagne (Rennes)	ENSA Bretagne (Rennes)	
		DIPAU	
		GRIEF Groupe de Recherche sur l'Invention et l'Évolution des Formes	S2PUR170012712
	ENSA Clermont-Ferrand	ENSA Clermont-Ferrand	
		GERPHAU Groupe d'Études et de Recherches Philosophie, Architecture, Urbain	
		Marges & Transformations	
	ENSA Grenoble	ENSA Grenoble	
		AE&CC Architecture, Environnement et Cultures Constructives	S2PUR160009539
		Architecture, Paysage, Montagne	S2PUR160010526
		CRATERRE Culture constructives et développement durable	
		CRESSON Centre de recherche sur l'espace sonore & l'environnement urbain	
	ENSA Grenoble	MHAevt Les métiers de l'histoire de l'architecture, édifices-villes-territoires	S2PUR160009756
		UMR 1563 AAU Ambiances, Architectures, Urbanités	
	ENSA Lille (architecture et paysage)	LaCTH Laboratoire Conception Territoire Histoire	S2PUR150008566
		ENSA Lille (architecture et paysage)	
	ENSA Lyon	ENSA Lyon	
		LAURE Lyon, Architecture Urbanisme recherche	
		MAP ARIA Applications et Recherches en Informatique pour l'Architecture	
		UMR 3495 MAP Modèles et simulations pour l'architecture et le patrimoine	
	ENSA Marne-la-Vallée	ENSA Marne-la-Vallée	
		OCS Observatoire de la condition suburbaine	
	ENSA Marseille-Luminy	ENSA Marseille-Luminy	
		DREAM Dép. de la Recherche doctorale de l'École d'Architecture de Marseille	
		GAMSAU Groupe de recherche pour l'Application des Méthodes Scientifiques à l'Architecture et à l'Urbanisme	
		INAMA - Investigations sur l'histoire et l'actualité des mutations architecturales	DER-PUR180014679
INSARTIS			
UR Projects		DER-PUR180014677	
ENSA Montpellier	ENSA Montpellier		
	ARTOPOS		

		GRSSLA	
		LIFAM Laboratoire Innovation Formes Architectures Milieux	S2PUR150008683
	ENSA Nancy	ENSA Nancy	
		LHAC Laboratoire d'histoire d'architecture contemporaine	DER-PUR180015289
		MAP CRAI Centre de Recherche en Architecture et Ingénierie	
	ENSA Nantes	ENSA Nantes	
		CERMA Centre de recherche méthodologique d'architecture	
		GERSA Groupe d'étude et de recherche scénologique en architecture	
		LAUA Laboratoire langages actions urbaines altérités	
		ATE Normandie	S2PUR170012839
	ENSA Paris-Belleville	ENSA Paris-Belleville	
		IPRAUS Institut Parisien de recherche : architecture, urbanistique et société	
		LAA Laboratoire Architecture Anthropologie	
	ENSA Paris-La Villette	ENSA Paris-La Villette	
		AHTTEP Architecture Histoire Technique Territoire Patrimoine	
		AMP Architecture Milieu Paysage	
		ARIAM LAREA Informatique Architecture et Modélisation - Laboratoire d'Architecturologie et de Recherches Epistémologiques sur l'Architecture	
		LET Laboratoire Espace Travail	
		MAP MAACC Modélisations pour l'Assistance à l'Activité Cognitive de la Conception	
	ENSA Paris-Malaquais	ENSA Paris-Malaquais	
		GSA Laboratoire géométrie structure architecture	S2UR140005731
		Lab ACS Architecture, culture société XIX <sup>ème</sup> -XXI <sup>ème</sup> siècles	
		LIAT Laboratoire Infrastructures, Architecture, Territoire	S2UR140005732
	ENSA Paris-Val-de-Seine	ENSA Paris-Val-de-Seine	
		CRH Centre de recherche sur l'habitat	
		EVCAU - Espace Virtuel de Conception Architecturale et Urbaine	S2UR140007422
	ENSA Saint-Etienne	ENSA Saint-Etienne	
		Mutations et pratiques	
		TRANSFORMATION(S)	S2PUR160011175
	ENSA Strasbourg	ENSA Strasbourg	
		AMUP Architecture, morphologie/morphogenèse urbaine et projet	DER-PUR180014191
	ENSA Toulouse	ENSA Toulouse	
		ASM Archéologie des sociétés méditerranéennes	
		GRECAU : Groupe de Recherche Environnement Conception Architecturale et Urbaine	
		GRECO Groupe de Recherche Environnement Conception	
		LRA Laboratoire de Recherche en Architecture	S2PUR160010857
	ENSA Versailles	ENSA Versailles	
		GRAI Groupe de recherche sur l'architecture et les infrastructures	
		LADRHAUS Lab. de recherche histoire architecturale et urbaine – sociétés	
		LéaV - Laboratoire de recherche de l'école nationale supérieure d'architecture de Versailles	S2PUR150008725

Établissement public national (EPN)	Bibliothèque Nationale de France	
	Centre des Monuments nationaux	
	Centre Georges Pompidou	
	Château de Versailles	
	Cité de la Céramique - Sèvres & Limoges	
	Cité de l'architecture et du patrimoine	
	Conservatoire national supérieur d'art dramatique Paris	
	Conservatoire national supérieur de musique et de danse Lyon	
	École de Chaillot (Cité de l'architecture et du patrimoine)	
	École de danse de l'Opéra de Paris Nanterre	
	École du Louvre	
	École nationale supérieure d'art de Bourges	
	École nationale supérieure d'art de Dijon	
	École nationale supérieure d'art de Limoges	
	École nationale supérieure d'art de Nancy	
	École nationale supérieure d'art de Paris-Cergy	
	École nationale supérieure d'arts Nice -Villa Arson	
	École nationale supérieure de création industrielle - Les Ateliers Paris	
	École nationale supérieure de la photographie Arles	
	École nationale supérieure des arts décoratifs Paris	
	École nationale supérieure des Beaux-Arts Paris	
	École sup. d'art dramatique du théâtre national de Strasbourg (TNS)	
	INA Sup Bry sur Marne	
	INP Institut national du Patrimoine	
	INRAP Institut national de recherche en archéologie préventive	
	Institut national d'histoire de l'art (Inha)	S2UR140007095
	La FEMIS - École nat. supérieure des métiers de l'image et du son Paris	
	MUCEM Musée des Civilisations de l'Europe et de la Méditerranée	
Musée Cluny / Musée national du Moyen-Age		
Musée d'Orsay		
Musée du Louvre		
Musée du quai Branly Jacques Chirac	S2PUR160011226	
Musée national de la Renaissance		
Musée national des arts asiatiques Guimet		
GIP	ARC Nucleart	
	CICRP Centre Interdisciplinaire Conservation & Restauration du Patrimoine	
	SCRIME	
SCN	Archives nationales	
	Archives nationales d'Outre-Mer	
	Archives nationales du monde du travail	

		C2RMF Centre de recherche et de restauration des musées de France	
		DRASSM Département des recherches archéologiques subaquatiques et sous-marines	
		LRMH Laboratoire de recherche des monuments historiques	
		Manufacture de Beauvais	
		Manufacture de la Savonnerie	
		Médiathèque de l'architecture et du patrimoine	
		Musée d'archéologie national	
UMR Unités mixtes de recherche	MC	CMBV Centre de musique baroque de Versailles	
	MC / CNRS / U Bordeaux 3	FR 3383 FSAB Fédération des sciences archéologiques de Bordeaux	S2SF160009983
	MC	FR 3506 NEW AGLAE L'accélérateur Grand Louvre d'analyses élémentaires	
	MC / CNRS / UPMC	LAM Lutheries Acoustique Musique	S2UR140005424
	CNRS / EHESS (MC)	LAHIC Labo. d'Anthropol. et d'Histoire de l'Institution de la Culture	S2UR140005313
	CNRS / ENSA Grenoble/ ENSA Nantes/ École Centrale Nantes	UMR 1563 AAU Ambiances, Architectures, Urbanités	S2PUR160009656
	MC / CNRS / ENSA Paris-Belleville / ENSA Paris-Malaquais / ENSA Paris-La Villette / ENSA Marne la Vallée	UMR 3329 AUSSer Architecture, Urbanisme, Société : savoirs, enseignements, recherche	S2UR140005544
	MC / CNRS / ENSA Lyon / ENSA Nancy / ENSA Marseille / ENSA Paris La Villette	UMR 3495 MAP Modèles et simulations pour l'architecture et le patrimoine	DER-PUR180014565
	MC / CNRS / INRAP / U Montpellier 3	UMR 5140 ASM Archéologie des sociétés méditerranéennes	S2PUR150008257
	MC / CNRS/ U. Bordeaux (INRAP)	UMR 5199 PACEA De la Préhistoire à l'Actuel : culture, environnement et anthropologie	S2PUR160009578
	CNRS / U Bordeaux / U Pau / ENSAP Bordeaux	UMR 5319 PASSAGES	S2PUR160010611
	CNRS / U Lyon / Mines St Etienne / INSA Lyon / ENS Lyon / ENTPE / ENSA Lyon	UMR 5600 EVS Environnement, Ville, Société	S2PUR160010463
	MC / CNRS / U Toulouse 2 (INRAP, EHESS)	UMR 5608 TRACES Travaux et recherches archéologiques sur les cultures, les espaces et les sociétés	S2PUR160010973
	MC / CNRS / U Dijon / INRAP	UMR 6298 ArTeHis Archéologie, Terre, Histoire, Sociétés	S2PUR170011818
	MC / CNRS / U Rennes 1 / U Rennes 2 / U Nantes / U le Mans (INRAP)	UMR 6566 CREAAH Centre de recherche en archéologie, archéosciences, histoire	S2PUR170011827
	MC / CNRS / U Paris Nanterre / INRAP / U Panthéon Sorbonne	UMR 7041 ArScAn Archéologies et sciences de l'antiquité	S2UR140005500
MC / CNRS / U Haute Alsace / U Strasbourg (INRAP)	UMR 7044 ARCHIMEDE Archéologie et histoire ancienne : Méditerranée et Europe	DER-PUR180014541	
MC / CNRS / U Paris Ouest Nanterre / U Paris8 / ENSA Paris-	UMR 7218 LAVUE Laboratoire Architecture, Ville, Urbanisme, Environnement	S2UR140007148	

La Villette / ENSA P.-Val de Seine		
MC / CNRS / U Aix-Marseille	UMR 7269 LAMPEA Laboratoire méditerranéen de Préhistoire Europe-Afrique	DER-PUR180014892
CNRS / U Orléans / U Tours / BNF	UMR 7270 LLL Laboratoire ligérien de linguistique	DER-PUR180014569
MC / CNRS / U Poitiers	UMR 7302 CESC Centre d'études supérieures de civilisation médiévale	DER-PUR180015205
MC / CNRS / U Tours	UMR 7323 CESR Centre d'études supérieures de la Renaissance	DER-PUR180013954
MC / CNRS / U Paris Sorbonne	UMR 8150 Centre André Chastel	S2UR140006587
MC / CNRS / U Lille 3	UMR 8164 HALMA Histoire, archéologie, littérature des mondes anciens	S2PUR150008215
MC / CNRS / U Panthéon Sorbonne	UMR 8218 ACTE Arts, Créations, Théories, Esthétiques	S2UR140005502
MC / CNRS / BNF / U Paris Sorbonne	UMR 8223 IReMus Institut de recherche en musicologie	S2UR140005557
MC / CNRS / UPMC / Ircam	UMR 9912 STMS Sciences et technologies de la musique et du son	S2UR140005463
CNRS / INHA	USR 3103 IN VISU - L'information visuelle et textuelle en histoire de l'art : nouveau terrains, corpus, outils	S2UR140005548
MC / CNRS / MNHN	USR 3224 CRC Centre de recherche sur la conservation	S2UR140005308
MC / CNRS / U Versailles Saint-Quentin	USR 3461 IPANEMA Institut photonique d'analyse non destructive européen des matériaux anciens	S2PUR150008031

## Annexe II - Table de correspondance « Types de doc /Production » entre sources

Catégorie générique	HCÉRES Code	HCÉRES description	HAL code	HAL catégorie	HAL description	WoS Code	WoS description
ARTICLE	ASCL	Articles dans des revues sans comité de lecture.	ART	Publication	Article dans une revue		
	ACLN	Articles dans des revues avec comité de lecture non répertoriées dans des bases de données internationales ou par le Hcéres					
	ACL	Articles dans des revues internationales ou nationales avec comité de lecture répertoriées par l'HCÉRES ou dans les bases de données internationales (ISI Web of Knowledge, Pub Med...).					
						Article	Reports of research on original works. Includes research papers, features, brief communications, case reports, technical notes, chronology, and full papers that were published in a journal and/or presented at a symposium or conference.
						Correction	Correction of errors found in articles that were previously published and which have been made known after that article was published. Includes additions, errata, and retractions.
						Reprint	An article that was previously published.
						Review	A renewed study of material previously studied. Includes review articles and surveys of previously published literature. Usually will not present any new information on a subject.
						Editorial Material	An article that gives the opinions of a person, group, or organization. Includes editorials, interviews, commentary, and discussions between individual, post-paper discussions, round table symposia, and clinical conferences.
						Letter	Contributions or correspondence from the readers to the journal editor concerning previously published material.
	ACTI	Communications avec actes dans un congrès international.	COMM/ART	Publication	Communication dans un congrès ayant un DOI	Proceedings Paper	Published literature of conferences, symposia, seminars, colloquia, workshops, and conventions in a wide range of disciplines. Generally published in a book of conference proceedings.
	ACTN	Communications avec actes dans un congrès national.					
COMMUNICATION	COM	Communications orales sans actes dans un congrès international ou national.	COMM	Publication	Communication dans un congrès	Meeting Abstract	A general summation of completed papers that were or will be presented at a symposium or conference.

	<b>AFF</b>	Communications par affiche dans un congrès international ou national.	<b>POSTER</b>	Publication	Poster		
	<b>INV</b>	Conférences données à l'invitation du Comité d'organisation dans un congrès national ou international.					
OUVRAGE	<b>OS</b>	Ouvrages scientifiques	<b>OUV</b>	Publication	Ouvrage (y compris édition critique et traduction)		
	<b>DO</b>	Directions d'ouvrages ou de revues.	<b>DOUV</b>	Publication	Direction d'ouvrage, Proceedings, Dossier		
	<b>OS-chap</b>	Chapitres dans ouvrages scientifiques	<b>COUV</b>	Publication	Chapitre d'ouvrage	<b>Book Chapter</b>	A monograph or publication written on a specific topic within a main division in a book.
Autre produit	<b>OV</b>	Ouvrages de vulgarisation					
	<b>OV-chap</b>	Chapitres dans ouvrages de vulgarisation					
	<b>AP</b>	Autres productions: bases de données, logiciels enregistrés, traductions, comptes rendus d'ouvrages, rapports de fouilles, guides techniques, catalogues d'exposition, rapports intermédiaires de grands projets internationaux, etc.	<b>PATENT</b>	Publication	Brevet		
			<b>OTHER</b>	Publication	Autre publication		
			<b>THESE</b>	Doc non publié	Thèse		
			<b>OTHERREPORT</b>	Doc non publié	Autre rapport, séminaire, workshop		
			<b>PRESCONF</b>	Doc non publié	Document associé à des manifestations scientifiques		
			<b>REPORT</b>	Doc non publié	Rapport		
			<b>UNDEFINED</b>	Doc non publié	Pré-publication, Document de travail		
			<b>HDR</b>	Travaux universitaires	Habilitation à diriger des recherches		
			<b>LECTURE</b>	Travaux universitaires	Cours		
			<b>MEM</b>	Travaux universitaires	Mémoire d'étudiant		
			<b>IMG</b>	Données de recherche	Image		
			<b>SON</b>	Données de recherche	Son		
			<b>VIDEO</b>	Données de recherche	Vidéo		
			<b>Art Exhibit Review</b>	Reviews of gallery or museum showings of artworks.			
			<b>Biographical-</b>	Obituaries, articles focusing on the life of an			



## Annexe III - Nomenclature disciplinaire

La nomenclature de l'OST en 11 grandes disciplines est définie par agrégation des spécialités scientifiques de Clarivate Analytics pour les sciences de la matière, de la vie, humaines et sociales dans le Web of Science®.

Le tableau suivant fournit la correspondance entre une grande discipline et les spécialités scientifiques qu'elle agrège. Dans la base, les revues peuvent être rattachées à plusieurs spécialités.

Grandes disciplines	Libellés anglais des spécialités dans le WoS	
<b>BIOLOGIE FONDAMENTALE</b>	ANATOMY & MORPHOLOGY BEHAVIORAL SCIENCES BIOCHEMICAL RESEARCH METHODS BIOCHEMISTRY & MOLECULAR BIOLOGY BIOPHYSICS BIOTECHNOLOGY & APPLIED MICROBIOLOGY CELL & TISSUE ENGINEERING CELL BIOLOGY DEVELOPMENTAL BIOLOGY ENGINEERING, BIOMEDICAL EVOLUTIONARY BIOLOGY GENETICS & HEREDITY MATERIALS SCIENCE, BIOMATERIALS	MATHEMATICAL & COMPUTATIONAL BIOLOGY MEDICAL LABORATORY TECHNOLOGY MICROBIOLOGY MICROSCOPY NEUROIMAGING NEUROSCIENCES NUTRITION & DIETETICS PARASITOLOGY PHYSIOLOGY PSYCHOLOGY REPRODUCTIVE BIOLOGY VIROLOGY
<b>RECHERCHE MEDICALE</b>	ALLERGY ANDROLOGY ANESTHESIOLOGY AUDIOLOGY & SPEAK-LANGUAGE PATHOLOGY CARDIAC & CARDIOVASCULAR SYSTEMS CHEMISTRY, MEDICINAL CLINICAL NEUROLOGY CRITICAL CARE MEDICINE DENTISTRY, ORAL SURGERY & MEDICINE DERMATOLOGY EMERGENCY MEDICINE ENDOCRINOLOGY & METABOLISM GASTROENTEROLOGY & HEPATOLOGY GERIATRICS & GERONTOLOGY HEALTH CARE SCIENCES & SERVICES HEMATOLOGY IMMUNOLOGY INFECTIOUS DISEASES INTEGRATIVE & COMPLEMENTARY MEDICINE MEDICAL ETHICS MEDICINE, GENERAL & INTERNAL MEDICINE, LEGAL MEDICINE, RESEARCH & EXPERIMENTAL NURSING OBSTETRICS & GYNECOLOGY ONCOLOGY	OPHTHALMOLOGY ORTHOPEDICS OTORHINOLARYNGOLOGY PATHOLOGY PEDIATRICS PERIPHERAL VASCULAR DISEASE PHARMACOLOGY & PHARMACY PRIMARY HEALTH CARE PSYCHIATRY PUBLIC, ENVIRONMENTAL & OCCUPATIONAL HEALTH RADIOLOGY, NUCLEAR MEDICINE & MEDICAL IMAGING REHABILITATION RESPIRATORY SYSTEM RHEUMATOLOGY SPORT SCIENCES SUBSTANCE ABUSE SURGERY TOXICOLOGY TRANSPLANTATION TROPICAL MEDICINE UROLOGY & NEPHROLOGY VETERINARY SCIENCES
<b>BIOLOGIE APPLIQUEE - ECOLOGIE</b>	AGRICULTURAL ECONOMICS & POLICY AGRICULTURAL ENGINEERING AGRICULTURE, DAIRY & ANIMAL SCIENCE AGRICULTURE, MULTIDISCIPLINARY AGRONOMY BIODIVERSITY CONSERVATION BIOLOGY ECOLOGY ENTOMOLOGY FISHERIES	FOOD SCIENCE & TECHNOLOGY FORESTRY HORTICULTURE MATERIALS SCIENCE, TEXTILES MYCOLOGY ORNITHOLOGY PLANT SCIENCES SOIL SCIENCE ZOOLOGY
<b>CHIMIE</b>	CHEMISTRY, ANALYTICAL CHEMISTRY, APPLIED CHEMISTRY, INORGANIC & NUCLEAR CHEMISTRY, MULTIDISCIPLINARY CHEMISTRY, ORGANIC CHEMISTRY, PHYSICAL CRYSTALLOGRAPHY ELECTROCHEMISTRY	MATERIALS SCIENCE, CERAMICS MATERIALS SCIENCE, CHARACTERIZATION & TESTING MATERIALS SCIENCE, COATINGS & FILMS MATERIALS SCIENCE, COMPOSITES MATERIALS SCIENCE, MULTIDISCIPLINARY MATERIALS SCIENCE, PAPER & WOOD NANOSCIENCE & NANOTECHNOLOGY POLYMER SCIENCE
<b>PHYSIQUE</b>	ACOUSTICS INSTRUMENTS & INSTRUMENTATION OPTICS PHYSICS, APPLIED PHYSICS, ATOMIC, MOLECULAR & CHEMICAL PHYSICS, CONDENSED MATTER	PHYSICS, FLUIDS & PLASMAS PHYSICS, MATHEMATICAL PHYSICS, MULTIDISCIPLINARY PHYSICS, NUCLEAR PHYSICS, PARTICLES & FIELDS SPECTROSCOPY
<b>SCIENCES DE L'UNIVERS</b>	ASTRONOMY & ASTROPHYSICS ENGINEERING, ENVIRONMENTAL ENGINEERING, GEOLOGICAL ENVIRONMENTAL SCIENCES	LIMNOLOGY MARINE & FRESHWATER BIOLOGY METEOROLOGY & ATMOSPHERIC SCIENCES MINERALOGY

	GEOCHEMISTRY & GEOPHYSICS GEOGRAPHY, PHYSICAL GEOLOGY GEOSCIENCES, MULTIDISCIPLINARY	OCEANOGRAPHY PALEONTOLOGY WATER RESOURCES
<b>SCIENCES POUR L'INGENIEUR</b>	AUTOMATION & CONTROL SYSTEMS CONSTRUCTION & BUILDING TECHNOLOGY ENERGY & FUELS ENGINEERING, AEROSPACE ENGINEERING, CHEMICAL ENGINEERING, CIVIL ENGINEERING, ELECTRICAL & ELECTRONIC ENGINEERING, INDUSTRIAL ENGINEERING, MANUFACTURING ENGINEERING, MARINE ENGINEERING, MECHANICAL ENGINEERING, MULTIDISCIPLINARY	ENGINEERING, OCEAN ENGINEERING, PETROLEUM IMAGING SCIENCE & PHOTOGRAPHIC TECHNOLOGY MECHANICS METALLURGY & METALLURGICAL ENGINEERING MINING & MINERAL PROCESSING NUCLEAR SCIENCE & TECHNOLOGY OPERATIONS RESEARCH & MANAGEMENT SCIENCE REMOTE SENSING THERMODYNAMICS TRANSPORTATION SCIENCE & TECHNOLOGY
<b>INFORMATIQUE</b>	COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE COMPUTER SCIENCE, CYBERNETICS COMPUTER SCIENCE, HARDWARE & ARCHITECTURE COMPUTER SCIENCE, INFORMATION SYSTEMS COMPUTER SCIENCE, INTERDISCIPLINARY APPLICATIONS COMPUTER SCIENCE, SOFTWARE ENGINEERING	COMPUTER SCIENCE, THEORY & METHODS MEDICAL INFORMATICS LOGIC ROBOTICS TELECOMMUNICATIONS
<b>MATHEMATIQUES</b>	MATHEMATICS MATHEMATICS, APPLIED	MATHEMATICS, INTERDISCIPLINARY APPLICATIONS STATISTICS & PROBABILITY
<b>SCIENCES HUMAINES</b>	ANTHROPOLOGY ARCHAEOLOGY ARCHITECTURE AREA STUDIES ART ASIAN STUDIES CLASSICS COMMUNICATION DANCE DEMOGRAPHY ETHICS ETHNIC STUDIES FILM, RADIO, TELEVISION FOLKLORE GEOGRAPHY HISTORY HISTORY & PHILOSOPHY OF SCIENCE HISTORY OF SOCIAL SCIENCES HUMANITIES, MULTIDISCIPLINARY LANGUAGE & LINGUISTICS LINGUISTICS LITERARY REVIEWS LITERARY THEORY & CRITICISM LITERATURE	LITERATURE, AFRICAN, AUSTRALIAN, CANADIAN LITERATURE, AMERICAN LITERATURE, BRITISH ISLES LITERATURE, GERMAN, DUTCH, SCANDINAVIAN LITERATURE, ROMANCE LITERATURE, SLAVIC MIDDLE & RENAISSANCE STUDIES MUSIC PHILOSOPHY POETRY PSYCHIATRY PSYCHOLOGY, APPLIED PSYCHOLOGY, BIOLOGICAL PSYCHOLOGY, CLINICAL PSYCHOLOGY, DEVELOPMENTAL PSYCHOLOGY, EDUCATIONAL PSYCHOLOGY, EXPERIMENTAL PSYCHOLOGY, MATHEMATICAL PSYCHOLOGY, MULTIDISCIPLINARY PSYCHOLOGY, PSYCHOANALYSIS PSYCHOLOGY, SOCIAL RELIGION THEATER
<b>SCIENCES SOCIALES</b>	BUSINESS BUSINESS, FINANCE CRIMINOLOGY & PENOLOGY CULTURAL STUDIES ECONOMICS EDUCATION & EDUCATIONAL RESEARCH EDUCATION, SPECIAL ENVIRONMENTAL STUDIES ERGONOMICS FAMILY STUDIES GERONTOLOGY HEALTH POLICY & SERVICES HOSPITALITY, LEISURE, SPORT & TOURISM INDUSTRIAL RELATIONS & LABOR INFORMATION SCIENCE & LIBRARY SCIENCE INTERNATIONAL RELATIONS LAW	MANAGEMENT NURSING PLANNING & DEVELOPMENT POLITICAL SCIENCE PUBLIC ADMINISTRATION PUBLIC, ENVIRONMENTAL & OCCUPATIONAL HEALTH REHABILITATION SOCIAL ISSUES SOCIAL SCIENCES, BIOMEDICAL SOCIAL SCIENCES, INTERDISCIPLINARY SOCIAL SCIENCES, MATHEMATICAL METHODS SOCIAL WORK SOCIOLOGY SUBSTANCE ABUSE TRANSPORTATION URBAN STUDIES WOMEN'S STUDIES
<b>CATEGORIE MULTIDISC.</b>	EDUCATION, SCIENTIFIC DISCIPLINES	MULTIDISCIPLINARY SCIENCES

Source : <https://clarivate.com/products/web-of-science/>

## Annexe IV – Indice d'impact

Indice d'impact le corpus MC-WoS 10 et pour le total des publications de la France, 2008-13

Domaine de recherche	Publications MC-WoS 10		3. <i>Impact du total des publications France par rapport au monde</i>
	1. Impact par rapport aux publications de la France	2. Impact par rapport aux publications du monde	
<b>Anthropologie</b>	1,63	3,28	1,53
<b>Archéologie</b>	1,35	2,41	1,48
<b>Architecture</b>	1,36	0,96	0,72
<b>Chimie analytique</b>	1,15	1,08	0,99
<b>Expr. artistique, histoire de l'art, muséographie</b>	1,07	0,68	0,55
<b>Géographie physique</b>	0,59	0,86	1,39
<b>Géosciences</b>	0,6	0,87	1,37
<b>Histoire</b>	0,74	0,69	0,46
<b>Paléontologie</b>	0,7	0,93	1,11
<b>Sc. de l'environnement</b>	0,52	0,69	1,32

Source : Clarivate Analytics, traitements OST

Un indice supérieur à 1 signifie que les publications du périmètre considéré ont un nombre de citations moyen supérieur au nombre de citation par publication de la référence.

## **PARTIE II - Dispositif de suivi présenté le 6 avril 2018 au MC**

# CARACTÉRISATION DE LA RECHERCHE CULTURELLE

## — DISPOSITIF DE SUIVI DE LA PRODUCTION SCIENTIFIQUE DU MINISTÈRE DE LA CULTURE ATELIER DE CONCEPTION

A. de la Laurencie  
A. Glanard

6 avril 2018



## 1- PARTICIPATION À IPERU

### IPERU : INDICATEUR DE PRODUCTION DES ÉTABLISSEMENTS DE RECHERCHE UNIVERSITAIRE

- CARACTÉRISER LA PRODUCTION SCIENTIFIQUE ET TECHNOLOGIQUE DE CHAQUE ÉTABLISSEMENT ;
- SUIVRE L'ÉVOLUTION AU COURS DU TEMPS DE CETTE PRODUCTION ;
- POSITIONNER LA PRODUCTION DE L'ÉTABLISSEMENT DANS DES ESPACES TERRITORIAUX DE RÉFÉRENCE.

## 1- PARTICIPATION À IPERU

- L'INTERFACE DE REPÉRAGE DES PUBLICATIONS
- UN ATELIER PRATIQUE D'UTILISATION DE L'INTERFACE DE REPÉRAGE DES PUBLICATIONS SCIENTIFIQUES À DESTINATION DES CORRESPONDANTS
- LIVRAISON DES RAPPORTS DÉCRIVANT LEURS CARACTÉRISTIQUES BIBLIOMÉTRIQUES
- SERVICE PAYANT : ENTRE 2,3 K€ ET 3,2 K€



## 1- PARTICIPATION À IPERU

### PARTICIPER À IPERU POUR REPÉRER TOUS LES ANS LES PUBLICATIONS DU MIC

- Désigner un correspondant qui sera formé à l'utilisation de la plateforme
- Utiliser la liste des adresses envoyée par l'OST pour confirmer si une publication appartient au MIC ou non
- Enrichir cette liste



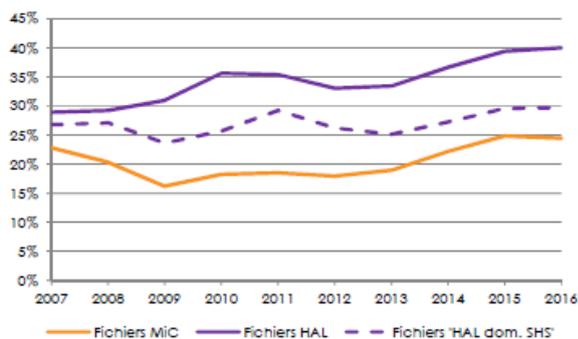
## 2- PORTAIL MIC SUR HAL

**CRÉATION D'UN PORTAIL MIC SUR HAL OÙ TOUTE LA PRODUCTION SERA DÉPOSÉE.**

**UTILISATION DU CORPUS CONSTITUÉ PAR L'OST**



## HAL MIC



Source : HAL, traitements OST du Hcéres

Alors que, dans l'ensemble de HAL la proportion de fichier déposés passe de 29 % à 40 % sur toute la période, pour le MIC les dépôts n'atteignent pas les 25 %.



## 2- PORTAIL MIC SUR HAL

- **Désigner un responsable de la plateforme qui sera formé à l'utilisation de HAL (Paris 07 ou champions HAL)**
  - Alimenter le portail du MIC
  - Administrer le portail
  - Créer et animer un réseau de correspondants dans les entités de recherche MIC
  - Constituer une véritable archive avec des PJ ou liens vers les documents
- **Avoir une politique d'incitation pour déposer les publications dans HAL**
- **Communiquer autour du portail MIC sur HAL**



### Portail : HAL INRAP



**Inrap - Institut national de recherches archéologiques préventives**  
 Date d'ouverture : 24 janvier 2010  
<https://hal-inrap.archives-ouvertes.fr>  
 Documents : 1 215 | Notices : 4 675 | Annexes : 18 | Dépôts dans le portail : 211

<https://hal-inrap.archives-ouvertes.fr/>

### Collection : ArTeHis UMR



**UMR 6298 ARTEHIS**  
 ARCHEOLOGIE  
 TEMPS  
 HISTOIRE  
 SOCIÉTÉS

**COLLECTION HAL D'ARTEHIS**

Accueil Consultation Rechercher Par axe (2017-2022) Ajouter le texte intégral

<https://hal-univ-bourgogne.archives-ouvertes.fr/ARTEHIS>



### 3- ÉTUDE RÉCURRENTE PAR L'OST

#### COMMANDE ANNUELLE OU BIANNUELLE À L'OST:

➤ METTRE À JOUR LE CORPUS DU MIC DANS LE WOS ET DANS HAL SUR LA BASE D'UNE LISTE FOURNIE PAR LE MIC

➤ CARACTÉRISER LA PRODUCTION DU MIC ET CALCULER LES INDICATEURS

➤ COÛT : 20K€ ANNUELS



