Rapport - n° 2017-A01 ▸ Mars 2017

Inspection générale des bibliothèques

# Les enjeux du numérique concernant les ressources documentaires en SHS

Rapport à

madame le ministre de l'éducation nationale,
de l'enseignement supérieur et de la recherche

monsieur le secrétaire d'Etat à l'enseignement
supérieur et à la recherche

# LISTE DES DESTINATAIRES

MADAME LA MINISTRE DE L'EDUCATION NATIONALE, DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE

M. OLIVIER NOBLECOURT, DIRECTEUR DE CABINET

MONSIEUR LE SECRETAIRE D'ETAT A L'ENSEIGNEMENT SUPERIEUR ET A LA RECHERCHE

M. CHRISTOPHE STRASSEL, DIRECTEUR DE CABINET

ENVOIS ULTÉRIEURS PROPOSÉS

- Madame la Directrice générale de l'enseignement supérieur et de l'insertion professionnelle
- Monsieur le directeur général pour la recherche et de l'innovation
- Monsieur le chef du service de la coordination des stratégies de l'enseignement supérieur et de la recherche
- Monsieur le sous-directeur du pilotage stratégique et des territoires
- Madame la cheffe du département de l'information scientifique et technique et du réseau documentaire
- Monsieur le président Du Haut-Conseil de l'évaluation de la recherche et de l'enseignement supérieur (HCERES)
- Monsieur le président de la conférence des présidents d'universités (CPU)
- Monsieur le président de la conférence des grandes écoles (CGE)
- Monsieur le président du comité de suivi de l'édition scientifique

**MINISTÈRE DE L'ÉDUCATION NATIONALE,**
**DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE**

———

*Inspection générale des bibliothèques*

# Les enjeux du numérique concernant les ressources documentaires en SHS

**Mars 2017**

Carole LETROUIT
*Inspecteur général des bibliothèques*

**SOMMAIRE**

**Introduction ..... - 1 -**

**1. A quel rythme la documentation acquise en SHS évolue-t-elle vers le numérique ? .. - 3 -**

1.1. Méthodologie ..... - 3 -

1.2. Résultats globaux ..... - 3 -

1.3. Répartition par support ..... - 5 -

1.4. Répartition par origine française ou étrangère pour les revues imprimées ..... - 6 -

1.5. Les livres électroniques ..... - 8 -

**2. Les freins à la transformation numérique des ressources documentaires en SHS - 11 -**

2.1. Manque de lisibilité de l'offre ..... - 11 -

*2.1.1. Offre éditoriale ..... - 11 -*

*2.1.2. Les particularités de l'édition française de recherche en SHS ..... - 12 -*

*2.1.3. Ressources numérisées ..... - 14 -*

2.2. Les archives scientifiques, des ressources documentaires en pleine mutation..... - 18 -

*2.2.1. La sauvegarde des archives sur support physique ..... - 18 -*

*2.2.2. Les archives scientifiques nativement numériques ..... - 19 -*

*2.2.3. Les archives d'Internet ..... - 21 -*

**3. Quels facteurs favorisent la transformation numérique des ressources documentaires en SHS ? ..... - 22 -**

3.1. Pourquoi la transformation numérique des ressources documentaires va s'accentuer ... - 22 -

*3.1.1. Le travail du chercheur a changé ..... - 22 -*

*3.1.2. La notion de publication évolue ..... - 22 -*

*3.1.3. Le travail du bibliothécaire a changé ..... - 23 -*

*3.1.4. Le rapport de la société à la science a évolué ..... - 23 -*

*3.1.5. L'édition scientifique change ..... - 24 -*

3.2. Comment favoriser la transformation numérique des ressources documentaires en SHS ? - 24 -

*3.2.1. Développer une offre de services autour des ressources documentaires à partir des besoins des chercheurs ..... - 24 -*

*3.2.2. Les services autour des données de la recherche* ..... *- 27 -*
*3.2.3. Les services autour des publications* ..... *- 29 -*
*3.2.4. La politique documentaire dans un contexte d'économie de services* ..... *- 33 -*
*3.2.5. Les services autour de l'édition* ..... *- 34 -*

**Conclusion** ..... **- 37 -**

**Préconisations** ..... **- 39 -**

**Annexes** ..... **- 41 -**

# Introduction

Le Plan d'action du secrétariat d'Etat chargé de l'Enseignement supérieur et de la recherche en faveur des sciences humaines et sociales comporte deux séries de mesures, les dix premières énoncées en juillet 2016 et cinq autres en novembre de la même année. La mesure 6 s'intitule « mieux intégrer les enjeux du numérique dans les formations et la recherche » et se décline en quatre actions dont une consiste à « missionner l'Inspection générale des bibliothèques pour réaliser un rapport sur les enjeux du numérique concernant les ressources documentaires en SHS (numérisation, indexation, nouveaux usages) ». Cette mission est inscrite au programme 2016-2017 de l'Inspection générale des bibliothèques.

En préambule à ce rapport, il convient de poser quelques définitions qui en délimiteront plus précisément les contours.

**Qu'entend-on ici par « SHS » ?**

Dans sa contribution au Plan pour les sciences humaines et sociales en date du 30 janvier 2016, l'Alliance ATHENA constate : « L'acronyme SHS pour Sciences Humaines et Sociales n'est guère compréhensible dans d'autres langues que le français et tend à laisser hors champ les arts et les lettres. L'expression « Humanités et sciences sociales » traduit mieux la réalité et les valeurs du domaine aujourd'hui. Les Humanités apparaissent, au plan mondial, comme une ressource morale et culturelle centrale ; les Sciences Sociales sont un outil d'analyse et de propositions indispensable à toute société démocratique ». Une fois l'acronyme transcrit dans une langue plus internationale, survient dans une logique académique la question des disciplines concernées. La France se distingue par une autre particularité : le périmètre des SHS n'est pas le même pour la formation et la recherche. Du point de vue de l'enseignement, le droit et les sciences économiques ne relèvent pas des SHS. Ces disciplines y sont en revanche intégrées quand il s'agit de nomenclatures et de programmes de recherche, parce que tel est le périmètre accepté à l'échelon international. De même, l'AERES devenue HCERES, regroupe en 15 listes les revues de recherche en SHS. Deux d'entre elles concernent le droit et l'économie-gestion. Seule la littérature n'apparaît pas.

Ce rapport ciblant en premier lieu les ressources documentaires pour la recherche, l'expression « SHS » y revêtira l'acception la plus large qui englobe les humanités (arts et lettres), les sciences humaines, les sciences sociales, le droit et les sciences économiques.

**Que sont les « ressources documentaires » pour les SHS ?**

Traditionnellement, les collections des bibliothèques universitaires sont constituées de livres et de périodiques, auxquels s'ajoutent en fonction des disciplines couvertes : des cartes, des partitions, des fonds iconographiques, des enregistrements audiovisuels, des archives, des manuscrits, des thèses…

Les débuts de l'ère numérique sont marqués par l'apparition des bases de données, des revues en ligne, des livres numérisés ou nativement numériques. S'y adjoignent maintenant des corpus numériques, des entrepôts d'archives … Cette diversification des ressources documentaires utilisées par la recherche sera embrassée autant que possible.

Il est fréquemment admis que les SHS accusent un retard dans le passage au numérique par rapport aux STM (sciences, techniques, médecine). Le présent rapport cherchera à établir dans une première partie si cette hypothèse se vérifie pour les ressources documentaires en étudiant à quel rythme la documentation acquise en SHS évolue vers le numérique. Il tentera ensuite d'identifier les facteurs qui freinent la transformation numérique de ces ressources documentaires. Cependant, cette évolution découle inévitablement de celle des modalités de la recherche : le travail du chercheur en SHS a profondément changé. Ce rapport envisagera enfin quelles synergies et quels développements de services autour des ressources documentaires peuvent contribuer à mieux intégrer les enjeux du numérique dans la recherche en SHS.

# 1. A quel rythme la documentation acquise en SHS évolue-t-elle vers le numérique ?

En 2009, le Groupement français de l'industrie et de l'information (GFII) a réalisé une étude sur « l'édition française en sciences sociales et humaines » qui comporte dans son volume 2 une analyse du marché français des ressources documentaires en SHS. Le marché institutionnel que constituent les bibliothèques universitaires est examiné de près à travers les données 2007 de l'Enquête statistique générale sur les bibliothèques universitaires (ESGBU). Une note méthodologique détaillée est jointe au rapport, ce qui permet de reproduire la démarche sur les données 2015 de l'ESGBU et d'établir une comparaison entre les deux séries de résultats.

## 1.1. Méthodologie

L'étude du GFII reprend la typologie établie dans ASIBU (Application statistique interactive des bibliothèques universitaires) qui distingue les universités à dominante « lettres et sciences sociales » de celles à dominante « droit et sciences économiques » ou encore des « petites universités pluridisciplinaires » et des « grandes universités pluridisciplinaires ». Les dépenses documentaires déclarées par les deux premières catégories d'université concernent directement les SHS. Cependant, il importe de prendre aussi en compte les dépenses documentaires en SHS des universités pluridisciplinaires. Comment les évaluer ? L'étude du GFII se base sur un ratio « dépenses documentaires par enseignant chercheur » pour estimer les dépenses documentaires en SHS des universités pluridisciplinaires à 20 % du total de leurs dépenses documentaires. Le montant global obtenu est ensuite augmenté de 18 % afin d'inclure les dépenses documentaires hors SCD, effectuées par d'autres services de l'université, en particulier des laboratoires ou unités de recherche. Ce pourcentage a été fixé à partir d'une enquête menée auprès des bibliothèques universitaires qui aboutit à une fourchette de 10 à 25 % de dépenses documentaires hors SCD en 2007. Le rapport de l'IGB sur les dépenses documentaires des universités remis en juin 2016 montre effectivement sur l'échantillon observé des amplitudes très variables (de 6 à 35 %) pour les dépenses documentaires hors SCD, tout en constatant un recul net entre 2011 et 2015. La valeur fixée par le GFII (18 %) sera néanmoins conservée à des fins de comparaison.

## 1.2. Résultats globaux

Ces principes méthodologiques étant posés, voici les résultats obtenus :

**Pour 2007,** l'étude du GFII établit les dépenses documentaires en SHS des universités françaises (SCD et hors SCD, c'est-à-dire autres services de l'université, dont unités de recherche) à **37,2 millions** d'euros qui représentent 35 % des dépenses documentaires des universités de tous types confondus. S'y ajoutent les dépenses documentaires des « grands

établissements », autre catégorie d'ASIBU, qui s'élèvent à **5,5 millions** d'euros. Entrent dans cette catégorie des institutions telles que la Fondation nationale des sciences politiques, la FMSH, l'INHA, le Collège de France, la Mazarine, le Museum d'histoire naturelle, l'EPHE, l'ENC, la bibliothèque Jacques Doucet, l'ENS Ulm…

Cette étude avance aussi un chiffrage des achats de ressources documentaires SHS des EPST sur la base d'une hypothèse selon laquelle la dépense moyenne annuelle en ressources documentaires pour un chercheur serait de 1000 euros, qui multipliés par un effectif de 2 535 chercheurs et assimilés en SHS donnent une somme de 2,5 millions. Ce calcul dont l'hypothèse de départ n'est pas étayée ne sera pas repris. L'Alliance Athena dénombrait 28 000 enseignants-chercheurs en SHS en 2012 dans les universités, grandes écoles, grands établissements et organismes de recherche dont près de 2000 au CNRS. Le rapport d'activité 2015 de ce dernier recense 1652 chercheurs à l'Institut des sciences humaines et sociales du CNRS (InSHS). Les dépenses documentaires sont morcelées entre les laboratoires locaux et les services centraux qui financent les abonnements nationaux, par exemple pour le portail BiblioSHS. De façon globale, le CNRS consacre chaque année environ 15 millions d'euros à l'achat/abonnement de ressources documentaires.

**Pour 2015**, l'application des mêmes principes aux données ESGBU donne un montant de **37,7 millions** d'euros pour les dépenses documentaires en SHS des universités françaises, qui représentent 33,12 % des dépenses documentaires des universités de tous types confondus (total ESGBU : 96,45 millions + 18 % = 113,81 millions), sans le coût de la licence nationale Elsevier et 28,93% de ces dernières si l'on ajoute aux 113,81 millions les 16,5 millions d'euros prélevés à la source sur le budget des universités au titre de la licence nationale Elsevier pour 2015[1].

Néanmoins, il a fallu introduire une étape supplémentaire afin de tenir compte de la fusion de plusieurs universités dans l'intervalle, ce qui déplace certaines des catégories à dominante « lettres et sciences sociales » ou « droit et sciences économiques » vers la catégorie « grandes pluridisciplinaires ». Il a été demandé aux directeurs des SCD fusionnés d'évaluer leurs dépenses documentaires en SHS.

Les dépenses des « grands établissements » s'élèvent à 5,9 millions d'euros.

Ces 37,7 millions se ventilent ainsi :

- Universités à dominante « lettres et sciences sociales » : 14 821 351 euros
- Universités à dominante « droit et sciences économiques » : 6 692 860
- Pluridisciplinaires (20 %) : 7 377 059
- Universités fusionnées (dépenses SHS) : 3 089 860 euros

Total : 31 622 275 euros (les dépenses telles que consignées dans l'ESGBU), auxquels on ajoute 18 % : **37 737 733 euros**

[1] Les établissements avaient reçu la consigne de ne pas inclure les montants Elsevier dans leurs calculs pour l'ESGBU.

La part des SHS dans les dépenses documentaires des universités n'a pas évolué sensiblement si l'on se limite aux données ESGBU (35 % en 2007, 33,12% en 2015). En revanche, elle diminue d'un peu plus de 6 % si l'on réintègre les coûts de la licence nationale Elsevier, alors que dans le même temps le volume global des dépenses des universités pour les ressources documentaires en SHS a légèrement augmenté (+ 500 000 euros, soit 1,3 %). Pour plus de précision encore, il faudrait ajouter la contribution des universités considérées à la licence nationale Elsevier. Elle est de moindre importance que celle des universités à dominante « science, technique, santé » et, surtout, elle peut difficilement être qualifiée de dépense pour des ressources en SHS, la part de ces disciplines étant très réduite dans la Freedom Collection d'Elsevier.

Ces chiffres reposent évidemment sur un choix de périmètre qui peut toujours prêter à discussion : les fusions d'établissements et l'apparition des COMUE dans l'ESGBU viennent perturber les classifications d'ASIBU. Les universités fusionnées ont bénéficié d'un traitement spécifique. Mais où ranger la COMUE Paris Sciences et Lettres (PSL) qui fait état d'une dépense de 949 832 euros pour la documentation électronique ? Le parti a été pris de ne pas l'inclure dans la catégorie à dominante « lettres et sciences sociales » afin de garder un ensemble d'établissements aussi semblable que possible à celui de 2007. Cependant, les dépenses de PSL incrémentent bien le montant total des dépenses documentaires établi à partir de l'ESGBU.

## 1.3. Répartition par support

**Pour 2007**, l'étude du GFII répartit ainsi les dépenses documentaires en SHS :

- Ouvrages imprimés : 40 %
- Revues imprimées : 38 %
- Documentation électronique : 22 %
-

Le traitement des **données 2015** aboutit à :

- Ouvrages imprimés : 35,54 %
- Revues imprimées : 21,21 %
- Documentation électronique : 42,21 %

Le delta de 1 % correspondant à l'acquisition d'autres supports (cd, dvd, cartes…)

Il est cependant intéressant d'examiner les variations qui se produisent d'une catégorie d'établissement à l'autre :

| 2015 | Dominante « Lettres, SHS » | Dominante « droit, sciences éco. » | Pluridisciplinaires | Fusionnées (part SHS) | moyenne |
|---|---|---|---|---|---|
| Livres imprimés | 39,5 % | 32,34 % | 27,33 % | 43 % | 35,54 % |
| Revues imprimées | 25 % | 21,28 % | 17,18 % | 21,4 % | 21,21 % |
| Documentation électronique | 33,4 % | 45,33 % | 54,53 % | 35,6 % | 42,21 % |
| autres | 2 % | 1 % | 1 % | 0 | 1 % |

Pour l'ensemble des établissements renseignant l'ESGBU, la répartition s'effectue ainsi :

- Ouvrages imprimés : 31,93 %
- Revues imprimées : 17 %
- Documentation électronique : 50 %

Le delta de 1,17 % correspond à l'acquisition d'autres supports (cd, dvd, cartes…)

Le profil SHS ne s'écarte donc que faiblement de la moyenne : 3, 5 % pour les ouvrages, 4 % pour les revues imprimées, moins de 8 % pour la documentation électronique.

## 1.4. Répartition par origine française ou étrangère pour les revues imprimées

L'étude du GFII procède ensuite à une analyse par origine française ou étrangère en se fondant toujours sur l'ESGBU. Entre temps cette enquête a été simplifiée et la distinction d'origine a disparu. Il a donc été demandé, à des fins de comparaison, aux directeurs de bibliothèques des universités à dominantes « Lettres, SHS » et « Droit, sciences économiques » de reprendre les données 2015 concernant les revues de façon à les scinder en « françaises » et « étrangères ». Les pluridisciplinaires, trop nombreuses, n'ont pas été sollicitées. Elles auraient dû d'abord isoler les ressources en SHS, ce qui supposait un travail considérable. Les universités fusionnées ont effectué cette double opération.

**En 2007**, la répartition était la suivante :

- Périodiques imprimés :

| | Français | étrangers |
|---|---|---|
| Lettres et SHS | 38 % | 62 % |
| Droit, sciences économiques | 46 % | 54 % |
| Pluridisciplinaires | 38 % | 62 % |
| Moyenne | 40 % | 60 % |

- Ressources électroniques :

| | Français | étrangers |
|---|---|---|
| Lettres et SHS | 35 % | 65 % |
| Droit, sciences économiques | 60 % | 40 % |
| Pluridisciplinaires | 35 % | 65 % |
| Moyenne | 43,3 % | 56,7 % |

**Pour 2015** :

- Périodiques imprimés :

| | Français | étrangers |
|---|---|---|
| Lettres et SHS | 31,2 % | 68,8 % |
| Droit, sciences économiques | 43,2 % | 56,8 % |
| Fusionnées | 38,2 % | 61,8 % |
| Moyenne | 37,5 % | 62,5 % |

- Revues numériques[2] :

| | Français | étrangers |
|---|---|---|
| Lettres et SHS | 27,2 % | 72,8 |
| Droit, sciences économiques | 46 % | 54 % |
| Fusionnées | 46,1 % | 53,9 % |
| Moyenne | 39,8 % | 60,2 % |

Le GFII s'est centré sur les informations financières qui lui permettaient de cerner le marché institutionnel. Il est possible de compléter l'analyse en mettant en regard le nombre d'abonnements auxquels correspondent les montants calculés. L'information est disponible pour 2007 sur le site ASIBU. La comparaison a été établie pour les seules universités à dominante « Lettres et SHS », cet ensemble ayant connu peu de variation de périmètre depuis 2007.

[2] La terminologie a évolué entre 2007 et 2015 : la notion de « base de données » n'a plus le même sens, désormais les bases de données comprennent quasiment toutes des documents en texte intégral (revues, livres, ...). Il vaut mieux se concentrer sur la rubrique « revues numériques » dont le contenu est plus précis. Autre indication terminologique : les adjectifs « électroniques » et « numériques » sont employés ici indifféremment. Il semble que le second tend à se substituer au premier sans évolution de sens notable.

| Bibliothèques | Nombre d'abonnements français 2007 | Nombre d'abonnements français 2015 | Nombre d'abonnements étrangers 2007 | Nombre d'abonnements étrangers 2015 |
|---|---|---|---|---|
| Bordeaux 3 | 735 | 574 | 691 | 537 |
| Grenoble 2 et 3 | 765 | 380 | 564 | 379 |
| Lille 3 | 666 | 412 | 664 | 384 |
| Lyon 2 | 879 | 449 | 568 | 272 |
| Lyon LSH/ENS Lyon | 141 | 466 | 344 | 288 |
| Nancy 2/lorraine SHS | 1275 | 820 | 764 | 273 |
| Rennes 2 | 971 | 472 | 647 | 303 |
| BNUS | 839 | 743 | 1267 | 1030 |
| Toulouse 2 | 722 | 869 | 972 | 368 |
| Paris 3 | 564 | 470 | 427 | 371 |
| Paris 4 | 505 | 276 | 602 | 378 |
| Paris 8 | 980 | 543 | 498 | 140 |
| Paris 10 | 986 | 572 | 769 | 252 |
| Sorbonne | 1422 | 826 | 3274 | 1903 |
| BSG | 1584 | 1475 | 687 | 565 |
| Sainte Barbe | 166 | 241 | 12 | 37 |
| BIULO/Bulac | 228 | 147 | 1065 | 793 |
| BDIC | 361 | 315 | 1233 | 787 |
| Total | 13 789 | 9230 | 15 048 | 8 787 |

Le nombre de titres français a diminué de 33,1 %, le nombre de titres étrangers de 41,6 %.

En 2007, la dépense des universités à dominante « Lettres, SHS » pour les périodiques imprimés s'élevait à 5 226 001 € dont 38 % pour des titres français (1 985 880 €) et par conséquent 62 % pour les titres étrangers (3 240 120 €).

En 2015, à l'intérieur d'une enveloppe qui a nettement diminué, la part des dépenses en titres français a régressé plus encore puisqu'elles ne forment plus que 31,2 % de l'ensemble des dépenses en périodiques imprimés pour les universités à dominante « Lettres, SHS », contre 38 % en 2007, et 43,2 % pour la dominante « droit, sciences économiques », alors que leur part atteignait 60 % en 2007. Les titres étrangers ont cependant fait l'objet de désabonnements plus massifs (-41,6 %) parce que soumis à des hausses de prix plus fortes.

## 1.5. Les livres électroniques

Au sein de la documentation électronique en SHS, le livre occupe une place qui varie du simple au double, tout restant globalement modeste :

- Dominante «Lettres, SHS » : les livres électroniques représentent 12,9 % des dépenses en documentation électronique
- Dominante « droit, sciences économiques » : 16 %
- Pluridisciplinaires : 8,6 %

Il est devenu très difficile de démêler les dépenses pour les revues électroniques de celles pour les bases de données. Ces dernières ne sont plus strictement bibliographiques et comprennent désormais très souvent des revues en texte intégral et même des livres qu'il est impossible de chiffrer séparément.

ASIBU qui s'est arrêté en 2010 n'identifie pas les livres numériques en tant que tels, mais distingue les abonnements électroniques des acquisitions de documents numériques. Etant donnée la diversité des modèles économiques, un livre numérique peut être inscrit dans l'une ou l'autre rubrique. Il est par conséquent impossible d'établir une comparaison. L'ESGBU distingue désormais d'un point de vue financier les livres électroniques acquis définitivement des livres électroniques soumis à abonnement.

Les bibliothèques universitaires à dominante SHS ont maintenu la masse financière consacrée aux livres imprimés et dégagé une nouvelle enveloppe pour les livres numériques qui a atteint 638 920 euros en 2015 (5 856 090 euros pour les livres imprimés). Ce constat est à rapprocher de deux autres phénomènes apparus depuis 2009, date à laquelle l'étude du GFII a été réalisée. Les plateformes initialement destinées à la diffusion de revues numériques en SHS se sont toutes ouvertes aux livres numériques : OpenEdition, Cairn, Persée et Jstor depuis peu. Le mouvement de l' « open access » axé naguère sur les revues s'étend désormais aux monographies. L'étude du GFII notait à ce propos : « les ouvrages – et non les seuls périodiques – restent dans le champ des SHS un support de communication prépondérant….En toute hypothèse, cette importance des achats d'ouvrages oblige à prendre en compte ce type de produit (encore très majoritairement diffusé sur support imprimé) dans les scénarios d'évolution possible des ressources documentaires SHS. Ce qui ne va pas sans difficultés : toutes les sources documentaires disponibles pour illustrer un éventuel passage à un modèle économique « open access » sont focalisées sur les seuls périodiques » (page 13, volume 2).

Même si la communication scientifique, au stade de la recherche, passe essentiellement par des articles dans des revues, parvenir à publier une monographie reste une apothéose pour les chercheurs en SHS.

Les caractéristiques des monographies n'ont pas empêché qu'elles migrent aussi dans l'environnement numérique, mais sans menacer de disparition les versions imprimées qui correspondent à des usages différents : la forme numérique sert essentiellement à faire des recherches, à sélectionner des passages intéressants, pas à une lecture continue. La problématique de l'accès libre s'est étendue aux livres, mais sans se focaliser sur deux options

(gold/green). Martin Paul EVE recense cinq modèles économiques différents[3], cette pluralité est sans doute liée à la montée en puissance récente de cette volonté de mettre également les livres en libre accès :

- Autoriser un dépôt en accès libre moyennant une période d'embargo et des « Book processing charges », comme pour les revues.
- Mettre en accès libre la version numérique « de base » et vendre des impressions à la demande et d'autres formats numériques (epub, PDF…). Open Book Publishers suit ce modèle, 92 titres ont été publiés, une plateforme de publication est proposée ainsi qu'une adhésion aux bibliothèques.
- Mettre en accès libre la version numérique « de base » et vendre des services et des formats numériques aux bibliothèques et aux particuliers (freemium)
- Mettre en place des subventions institutionnelles récurrentes et un service d'impression à la demande. Open Humanities Press suit ce modèle. L'université du Michigan et UCLA versent des financements.
- Recourir à un financement collectif : Knowledge Unlatched publie des listes d'ouvrages qui sont « libérés » si un nombre suffisant d'institutions contribuent. 449 livres en humanités et sciences sociales ont ainsi été mis en libre accès depuis 2014, mobilisant 270 bibliothèques et consortia de 21 pays.

Le DOAB (Directory of open access books) dénombre 5 922 livres académiques, certifiés par un comité de lecture, émanant de 169 éditeurs. Tous ne traitent pas des SHS, mais elles sont prépondérantes.

En 8 ans, le passage au numérique s'est accentué nettement puisque les dépenses en documentation électronique représentent désormais 42,21 %, contre 22 % en 2007, du montant total des dépenses documentaires en SHS. Ce dernier restant stable, on peut parler de déplacement des dépenses au détriment des revues imprimées qui perdent près de 17 % alors que les ouvrages imprimés n'accusent une baisse que de 4,5 % et peut-être 3,5 % si l'on suggère que le 1 % de dépense pour d'autres unités matérielles a été agrégé en 2007 aux dépenses en ouvrages, les marchés de monographies englobant parfois ces autres supports. Cette bonne résistance des ouvrages est à moduler en fonction de l'origine des titres. Le nombre de volumes acquis publiés en France a baissé de 9 % pour les universités à dominante « lettres, sciences sociales », entre 2007 et 2015 ; le nombre de volumes acquis publiés à l'étranger a chuté lui de 37,6 %. Il est donc très vraisemblable que le prix moyen d'un volume français ou étranger a augmenté dans l'intervalle.

**Le profil de dépenses pour les ressources documentaires en SHS s'est nettement rapproché du profil général puisque du côté SHS 42,21 % des dépenses sont dévolues aux ressources électroniques contre 50 % à l'échelon global.**

[3] Martin Paul EVE, "Open access and the Humanities ", Cambridge University Press, 2014, https://doi.org/10.1017/CBO9781316161012

# 2. Les freins à la transformation numérique des ressources documentaires en SHS

Il faut battre en brèche l'idée selon laquelle la documentation pour les SHS serait encore majoritairement sous forme imprimée et qu'il faut accélérer les processus de dématérialisation afin d'atteindre une couverture totale de l'imprimé par le numérique. D'une part, nous ne disposons pas d'une évaluation sérieuse du nombre de pages imprimées existantes, d'autre part nous ne savons pas précisément ce qui a déjà été numérisé[4] et surtout il importe de s'interroger d'abord sur les raisons qui justifieraient une dématérialisation totale. La Bibliothèque nationale de France (BNF) qui s'était lancée dans une campagne massive et systématique a vite compris que cette logique institutionnelle ne tenait pas sur la durée, principalement parce qu'elle nécessitait des moyens gigantesques, et lui a préféré une logique documentaire, fondée sur les besoins du public et des impératifs de préservation, comme l'explique le Bilan 2009-2014, achevé en décembre 2015, de la numérisation partenariale conduite par la BNF dans le cadre de sa politique de coopération. La BNF a donc défini des critères et des programmes de numérisation[5].

Afin de cerner les freins ou les facteurs favorisant la transformation numérique des ressources documentaires en SHS, il est indispensable de déterminer au préalable les objectifs de cette évolution qui ne peuvent se résumer à une volonté de dématérialisation totale et de dresser un état des lieux. Or, l'offre de documentation numérique en SHS manque de lisibilité.

## 2.1. Manque de lisibilité de l'offre

### *2.1.1. Offre éditoriale*

Qu'elles relèvent d'une offre éditoriale ou soient disponibles librement sur le web, ces ressources numériques se caractérisent par une dispersion extrême qui empêche d'en percevoir l'étendue et la diversité.

Concernant l'offre éditoriale, il n'existe aucun catalogue académique. Des plateformes agrègent des ressources au gré de leurs accords avec les éditeurs sans qu'aucune n'atteigne, loin s'en faut, l'exhaustivité. Le Consortium Couperin, par son travail de veille et de recensement, aide les établissements à compléter leur compréhension de l'offre éditoriale, mais l'éparpillement s'étale à nouveau lorsque ces mêmes établissements entreprennent de donner accès aux ressources souscrites ou achetées pour leur communauté académique : aucun logiciel ne parvient à fédérer la recherche sur toutes ces ressources documentaires réparties sur de multiples plateformes aux interfaces et aux structures de données différentes. Par conséquent, les usagers passent à côté de résultats pertinents pour leur recherche ou, dans le meilleur des cas, doivent recourir à un bibliothécaire pour saisir toutes les pièces de ce

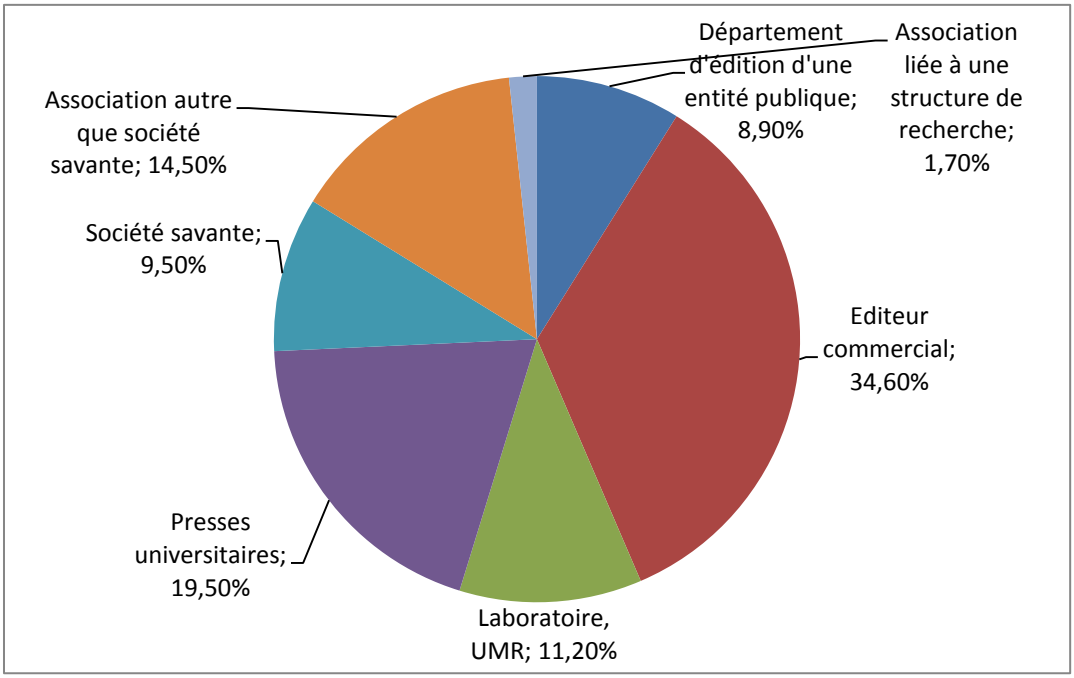
[4] Voir 2.1.2
[5] Voir 2.1.2

puzzle documentaire. Cette situation n'est pas propre aux SHS, mais s'y trouve aggravée par l'existence de ressources spécifiques non recensées dans les bases de connaissance des fournisseurs d'outils de découverte[6] ou ne respectant pas les standards internationaux.

### *2.1.2. Les particularités de l'édition française de recherche en SHS*

Au-delà du manque de lisibilité de l'offre éditoriale, on peut considérer que les particularités de l'édition française pour la recherche en SHS entravent la transformation numérique des revues et des livres publiés dans ce domaine. L'étude du GFII déjà mentionnée et plus récemment les travaux d'Odile Contat (InSHS – CNRS) ont décrit ce paysage éditorial et les fragilités qu'il génère.

« Dans le champ des SHS, la notion de communauté de recherche n'est pas déterminée a priori par la structuration internationale d'un champ de recherche disciplinaire ou sous-disciplinaire » (étude GFII, rapport de synthèse). D'autres mécanismes d'appartenance tels que l'école de pensée ou la tradition intellectuelle de l'organisme de rattachement jouent un rôle déterminant dans une sphère nationale centrée sur elle-même par la langue de publication utilisée : le français. Ces équipes de recherche s'expriment à travers une myriade de revues. On en dénombre plus d'un millier à parution régulière, le plus remarquable étant le nombre d'éditeurs qui les produisent puisque la majorité d'entre eux ne publient qu'une seule revue. Les abonnements par titre s'échelonnent entre 150 et 1200, la médiane s'établissant à 300 abonnements. Seules 2 % des revues SHS françaises sont publiées par des éditeurs privés internationaux. La répartition par type de structure éditrice illustre à merveille cet éparpillement :




Source : étude GFII, 2009

[6] Un outil de découverte est un logiciel permettant une recherche sur plusieurs ressources documentaires dont les données ont été rassemblées dans un index unique.

Cette dispersion a certainement retardé le passage au numérique, mais il a bien eu lieu. Les revues françaises en SHS ont opté massivement pour la mixité : elles ont maintenu le tirage papier et développé une version numérique dans l'espoir que les ventes de cette dernière compenseraient la baisse des abonnements à la version imprimée. 70 % des revues en SHS existeraient aujourd'hui sous les deux formes.

Le rapport de l'Institut des politiques publiques (IPP) sur « les revues en sciences humaines et sociales en France : libre accès et audience »[7], paru en juillet 2015, fournit une analyse détaillée de cette présence numérique des revues en SHS en s'appuyant sur des données 2014. Sur les 15 listes AERES en SHS, on recense 936 revues françaises dont 31,9 % sont présentes sur CAIRN ou Revues.org (page 44). Les trois disciplines les plus importantes en nombre de revues labellisées par l'AERES sont le droit (236 revues), l'histoire (179) et les arts (160). Ces trois disciplines sont faiblement représentées sur les deux plateformes mentionnées. Les mieux représentées sont : les sciences politiques (94 % des 48 revues de l'AERES sont sur CAIRN ou Revues.org), la géographie (75 % sur 48), la sociologie (85 % sur 69), les STAPS (95 % sur 43) et les sciences de l'information et de la communication (70 % sur 24). Cependant, sur les 637 revues non présentes sur ces deux plateformes, **seules 13 sont totalement absentes du web**. 66 % (419 revues) sont hébergées sur une autre plateforme, souvent Persée. En droit, 82 % des revues sont disponibles sur une autre plateforme, en histoire 76,4 %. L'existence d'un site web varie aussi beaucoup d'une discipline à l'autre.

Les chiffres du rapport de l'IPP induisent à conclure que la transformation numérique n'a que peu contribué à rassembler l'offre de revues. Il serait intéressant d'effectuer un nouveau pointage de la présence des revues labellisées sur les trois grandes plateformes françaises Persée, CAIRN et Revues.org.

Concernant le livre de recherche, la France se caractérise au regard des pays anglo-saxons par la faiblesse des presses universitaires. Ce sont elles qui éditent très largement en Grande-Bretagne et aux Etats-Unis les livres en SHS. En France, les éditeurs généralistes (La Découverte, Le Seuil, Gallimard…) assurent encore 20 à 25 % de la publication des titres très spécialisés (contre 50 % en 1974). Ils sont confrontés à un phénomène qui touche l'ensemble de l'édition, et pas seulement les SHS, en France et ailleurs : la baisse continue de la moyenne des tirages et des ventes par titre. Ces dernières s'étalent entre 300 et 1200 exemplaires et la valeur médiane s'établit à 450. Parallèlement, le nombre de titres publiés ne cesse d'augmenter. Les chercheurs sont soumis à une exigence de publication et leurs effectifs s'étant accru depuis les années soixante, ils se sont à la fois spécialisés et produisent beaucoup plus d'écrits dont les prérequis intellectuels réduisent le public potentiel. Certains éditeurs n'ont trouvé d'autre solution que d'exiger des auteurs une participation aux frais de publication. Ceux-ci se retournent alors vers leur institution de rattachement pour solliciter une aide à la publication qui s'apparente à des « books processing charges », sur le modèle des « articles processing charges » que connaissent bien les chercheurs en STM. Les

[7] Rédigé par Maya Bacache-Beauvallet, Françoise Benhamou et Marc Bourreau.

structures éditoriales actuelles ne peuvent concilier leur logique commerciale avec les attentes des chercheurs en matière de diffusion scientifique.

### *2.1.3. Ressources numérisées*

Les ressources numérisées par des bibliothèques ou des chercheurs souffrent du même manque de visibilité. Personne ne saurait dire ce qui a été numérisé en France. En premier lieu parce que beaucoup de corpus ont été dématérialisés et enregistrés sur un disque dur pour les besoins propres d'un chercheur qui ne se soucie pas de les mettre à disposition du grand nombre ou dans l'attente d'une autorisation de diffuser qui n'a pas été sollicitée au préalable. Ensuite parce que les outils destinés à ce recensement ont fait long feu ou ne sont que partiellement conçus dans cet objectif.

L'Agence bibliographique de l'enseignement supérieur (ABES) et le TGE Adonis ont fait développer et ont maintenu pendant six ans un inventaire en ligne des corpus numérisés et des projets de numérisation des établissements de l'ESR, appelé **Numes.** 205 notices de corpus furent créées de 2009 à 2014 par 72 institutions. Cependant, l'application fut victime d'un piratage en 2014 et ne s'en releva pas. Elle avait auparavant essuyé des critiques sur le niveau de signalement choisi : celui du fonds et pas du document, ce manque de précision ne permettait pas dans certains cas d'éviter des doublons de numérisation.

- **Le catalogue collectif *Patrimoine numérique* du Ministère de la Culture**

Le Ministère de la Culture, pour sa part, a développé au début des années 2000 un catalogue collectif des collections numérisées, dénommé « Patrimoine numérique ». Il vise à valoriser les collections patrimoniales et à en faciliter l'accès. Il s'articule avec le projet européen Michaël. 3262 collections relevant de 1043 institutions y sont signalées, parmi lesquelles 14 bibliothèques universitaires et 33 centres ou laboratoires de recherche. Il devrait être prochainement refondu de façon à intégrer le moteur Collections du Ministère de la Culture qui agrège 70 bases de données et alimente Europeana.

- **Calames**

Depuis 2008, Calames permet aux établissements de l'ESR de décrire leurs fonds de manuscrits et d'archives. Il s'est constitué à partir de la rétroconversion du Catalogue général des manuscrits des bibliothèques publiques de France (CGM) et du répertoire Palme (Répertoire de manuscrits littéraires français contemporains). Il repose sur un encodage en EAD, norme permettant de restituer l'organisation hiérarchisée d'un fonds et les interrelations entre les composants ainsi que de conserver le principe d'héritage des informations entre les niveaux. Une cinquantaine d'établissements utilisent cet outil développé par l'ABES. Cette dernière alloue aussi des aides financières pour des conversions rétrospectives d'inventaires en EAD. A ce jour, Calames compte 1130 inventaires analytiques qui se décomposent en 775 000 niveaux descriptifs. Parmi ceux-ci 45 000 sont de type « image fixe », 1 900 de type « enregistrement sonore » et 1 500 de type « images animées ».

Sur ces 1 130 inventaires, environ les deux tiers sont des fonds d'archives scientifiques, décrits par des bibliothèques d'institutions diverses, des services d'archives, des laboratoires, des MSH, ou des sociétés savantes. Le tiers restant correspond à des collections de manuscrits littéraires ou non.

Le service Calames Plus offre la possibilité d'associer des images à des notices descriptives. Seuls cinq établissements y recourent. Ceux qui disposent d'une bibliothèque numérique préfèrent enchâsser dans la notice un lien vers celle-ci. 237 inventaires sur les 1 130 contiennent au moins une telle mise en relation, ce qui représente environ 65 000 niveaux descriptifs pointant chacun vers un document numérisé. Calames est moissonnable en OAI-PMH.

Calames contribue à la visibilité des fonds numérisés puisqu'il en signale une partie. Cependant, son rôle dans ce domaine n'a jamais été clairement énoncé. Il mériterait de l'être dans le cadre de l'infrastructure de recherche Collex-Persée.

En effet, le Groupement d'intérêt scientifique (GIS) Collex-Persée en cours de constitution aura expressément pour mission de porter une politique nationale concertée en matière de numérisation. L'élaboration de cette politique requiert au préalable un état des lieux des opérations de numérisation conduites ces dix dernières années par des institutions relevant de l'ESR dans une perspective scientifique. L'infrastructure Collex-Persée desservant les chercheurs, il importe de centrer son action sur des programmes ou des stratégies scientifiques, à distinguer d'une approche purement patrimoniale. Cette politique de numérisation devra aussi clarifier le rôle dévolu aux différentes plateformes et catalogues inscrits dans son périmètre : Persée, Gallica, Calames, Huma-Num…

Nous ne savons pas ce qui est numérisé dans la sphère de l'ESR, mais nous disposons d'un chiffre qui peut susciter bien des réflexions : les bibliothèques indiquent depuis quelques années dans l'ESGBU les dépenses en numérisation imputées sur leur budget. Cette somme s'élevait en 2015 pour les seules bibliothèques en SHS (universités à dominante SHS, Grands Etablissements, universités à dominante droit, sciences économiques) à 850 877 € (et à 970 954 € pour l'ensemble des bibliothèques renseignant l'ESGBU). Sachant que les opérations de numérisation sont aussi parfois financées par les services de la recherche ou des laboratoires, les montants en jeu indiquent une activité loin d'être négligeable.

- **Les corpus de l'ESR intégrés à Gallica bénéficient, eux, d'une bonne visibilité.**

A partir de 2009, la BNF a réorienté sa politique partenariale et déployé une stratégie de coopération numérique fondée sur une logique documentaire et non plus institutionnelle. Cette dernière vise à constituer des corpus « les plus exhaustifs possibles répondant à la demande du public et à des nécessités de préservation du patrimoine, grâce à des programmes coopératifs de numérisation ouverts à toutes les bibliothèques quelles que soient leur taille et leur tutelle » (Bilan 2009-2014 de la numérisation partenariale conduite par la BNF dans le cadre de sa politique de coopération, décembre 2015). Cet objectif a été décliné en programmes de trois types :

-disciplinaires (sciences juridiques, arts, Guerre 1914-1918…),

-d'intérêt régional,

-visant à compléter des corpus existants (patrimoine équestre…).

Sur les 176 institutions ayant contribué à ces opérations de numérisation entre 2009 et 2014, 51 relèvent de l'ESR. Leurs collections ont pu être adjointes à Gallica selon trois modalités distinctes :

-intégration des documents physiques provenant de ces institutions dans les marchés de dématérialisation et ateliers de la BNF,

-intégration de fichiers numériques émanant de ces institutions,

-moissonnage des bibliothèques numériques de ces partenaires.

34 établissements de l'ESR sur les 51 partenaires ont reçu une subvention pour numériser des fonds en respectant les normes et conditions définies par la BNF. Leurs fichiers ont ensuite été incorporés à Gallica. Un même partenaire a pu participer à plusieurs opérations.

Dans le domaine des SHS, l'INHA, Cujas, le Museum et la BDIC ont obtenu un soutien compris entre 200 000 et 300 000 € chacun sur la période considérée, principalement pour leur participation aux programmes disciplinaires de numérisation : l'INHA pour l'art, Cujas pour les sciences juridiques et la BDIC pour la Guerre de 1914-1918, la presse d'émigration et la Seconde guerre mondiale.

Ces 51 partenaires de l'ESR ont bénéficié au total de 2 717 542 €, soit 44 % du montant total alloué par la BNF à la numérisation partenariale et produit 8 013 278 pages numériques, soit 46 % du volume global résultant de la numérisation partenariale. Cette implication s'explique par les relations qui existaient déjà entre un certain nombre de ces établissements et la BNF dont ils étaient pôles associés pour les acquisitions avant la réorientation de la politique partenariale de la BNF et par leur engagement antérieur dans des opérations de numérisation.

Toutefois ces corpus bien exposés dans Gallica ne représentent qu'une partie des dématérialisations effectuées ne serait-ce que pendant cette période 2009-2014.

- **Les appels à projets de numérisation BSN 5**

En 2013 et 2014, dans le cadre du segment 5 de la Bibliothèque scientifique numérique (BSN) pilotée par le MESR, ont été lancés deux appels à projets de numérisation. Les deux éditions remportèrent un vif succès auprès des établissements : 78 projets déposés en 2013 pour un montant total de 3,5 millions d'euros, 44 en 2014 pour 4 millions. La première année, 16 projets furent retenus pour 650 000 euros, la seconde 17 pour 646 000 euros. Une grande majorité des projets ressortaient des SHS ou de l'histoire des sciences. Le cahier des charges n'imposait pas de contraintes fortes en matière d'interopérabilité et de diffusion. La qualité technique de l'ensemble du projet et plus particulièrement la qualité d'indexation du corpus a été évaluée de même que la pertinence de la solution de mise en ligne, le caractère innovant du projet et l'archivage des données. Des institutions variées s'étaient mobilisées : universités, organismes de recherche, grandes écoles, laboratoires, UMR…Les bibliothèques portaient la

majorité des projets retenus, mais des laboratoires figuraient aussi parmi les lauréats. En 2014, trois organismes de recherche ont obtenu des financements importants pour leurs opérations de numérisation (Institut Pasteur : 79 000 €, IRD : 49 000, INRA Toulouse : 31 000).

Le MESR a ensuite décidé d'inclure la numérisation des collections dans le périmètre d'action de l'infrastructure de recherche Collex-Persée. Contrairement aux projets partenariaux sous l'égide de la BNF, ceux financés dans le cadre de BSN5 n'ont que peu de visibilité, si tant est qu'ils aient été menés à bien, ce qu'aucun document n'atteste.

- **Isidore**

Actuellement, l'outil qui déploie la plus grande efficacité pour la visibilité des ressources documentaires en SHS n'est pas un catalogue ou une bibliothèque numérique, mais un moteur de recherche dédié aux SHS : Isidore, développé par l'équipe de la Très Grande Infrastructure de Recherche (TGIR) Huma-Num. Il indexe aussi bien le texte intégral des documents que les métadonnées et il enrichit les données ainsi collectées à l'aide de référentiels sémantiques généralistes (Rameau, LCSH…) ou propres aux SHS (Gemet, GeoEthno, GeoNames, PACTOLS, index OpenEdition…). Il sait moissonner trois formats : OAI-PMH, RDF, flux Rss. Les 4,6 millions de documents moissonnés émanent de 4 200 sources et sont soumis à un million de concepts appartenant à des référentiels. Le passage à une indexation non seulement en français, mais aussi en anglais et en espagnol a nécessité un alignement des référentiels et donc une multiplication par trois des concepts. D'autres langues sont à l'étude puisque Isidore développe une politique d'internationalisation et connaît un franc succès en Amérique du Sud et du Nord. A ce jour, 10 % des sources sont anglo-saxonnes. Cette notion de source recouvre des réalités disparates : l'annuaire des sources recense aussi bien les Bibliothèques virtuelles humanistes que la Bibliothèque numérique du Limousin, la Bibliothèque municipale de Roubaix, le Centre de recherche bretonne et celtique, e-corpus, e-codices (bibliothèque virtuelles de manuscrits en Suisse), Calames, Cairn.info, le Centre Alexandre Koyré, HAL, Persée, les archives de données issues de la statistique publique, Fabula, Gallica (certaines parties), Erudit…et de nombreuses bibliothèques numériques de SCD ou de BIU. Les institutions qui souhaitent devenir une source d'Isidore, donc être moissonnées par ce dernier, doivent le faire savoir aux gestionnaires de cet outil. L'équipe se compose de 11 ingénieurs et chercheurs du CNRS regroupés dans une UMS (Unité mixte de service). Le budget de 1,4 millions d'euros provient de la DGRI du MESR.

Ce moteur de recherche très performant bénéficie peu de promotion de la part des bibliothécaires si l'on en croit les investigations menées sur les sites de plusieurs SCD et BIU. Il mériterait aussi d'être davantage mis en avant dans les formations à la recherche documentaire. Les établissements devraient en outre demander beaucoup plus systématiquement à être moissonnés lorsqu'ils ont mis en ligne des corpus numérisés dans le champ des SHS. En effet, même s'il englobe un nombre de documents impressionnant, tous les corpus qui pourraient entrer dans le périmètre d'Isidore n'y sont pas. Cette absence peut s'expliquer par un manque d'information sur l'existence de ce moteur de recherche, mais aussi par des prérequis techniques (hormis la compatibilité OAI-PMH), liés en particulier à la taille des fichiers, qu'il conviendrait d'expliciter. Il serait également utile de mesurer la

notoriété et l'usage du moteur de recherche Isidore, d'une part, des services Huma-Num d'autre part, auprès des chercheurs en SHS.

Isidore moissonne des contenus très disparates dont le dénominateur commun est d'appartenir au champ des SHS. S'y trouvent des articles, des bibliographies, des billets de blog, des expositions, des partitions, des cartes et plans, des recensions, des photos, des fonds d'archives… Ces contenus relèvent pour une part des « **archives scientifiques** », si l'on entend par là les productions des chercheurs ou des structures de recherche, et pour une autre part des « **matériaux de la recherche** » si l'on désigne ainsi ce qui est produit par des acteurs extérieurs à la recherche, mais collecté en vue de projets de recherche comme matériaux source (témoignages, photos, cartes…). Ce que l'on appelle désormais les « **données de la recherche** » constitue un sous-ensemble des archives scientifiques.

## 2.2. Les archives scientifiques, des ressources documentaires en pleine mutation

L'offre éditoriale en SHS ne constitue que la partie émergée de l'iceberg. Les chercheurs en SHS découvrent au gré de leurs travaux des gisements documentaires beaucoup plus vastes et invisibles parce que peu ou pas répertoriés. Les fonds numérisés affleurent par endroit. Les archives scientifiques forment des archipels sous-marins encore inexplorés pour la plupart.

### *2.2.1. La sauvegarde des archives sur support physique*

Les SHS se sont préoccupées plus tardivement que les sciences exactes de leurs archives parce qu'elles se trouvaient par ce biais confrontées à la question de leur légitimité scientifique : une réflexion sur l'organisation des archives implique un questionnement sur les conditions de validation des résultats et de réutilisation des données. Elle contraint une discipline à examiner ses présupposés et ses méthodes, sujets toujours sensibles en SHS. Un premier rapport confié à Roxane Silberman sur « les sciences sociales et leurs données » paraît en 1999. Ce sont finalement des Maisons des sciences de l'homme qui entreprennent au début des années 2000, avec le soutien de la direction des Archives de France et du CNRS, de sauvegarder ces fonds. En 2004, est lancé un programme national sur les archives de la recherche en sciences humaines et sociales (ARSHS), coordonné par la MSH de Dijon, en partenariat avec les MSH de Nanterre et Aix en Provence. Un questionnaire est adressé en 2005-2006 d'une part aux archives publiques afin d'identifier les fonds déjà préservés, et d'autre part à des laboratoires et unités de recherche afin de repérer les fonds ne bénéficiant d'aucun traitement. Les résultats sont enregistrés sous un logiciel qui produit des instruments de recherche archivistiques. Un catalogue collectif des fonds de la recherche en SHS voit ainsi le jour. Il porte le nom de « Portail des instruments de recherche archivistiques » et comprend

un portail « Archives numériques et données de la recherche en SHS » (PANDOR), l'ensemble étant géré par la MSH de Dijon.

Cette enquête a suscité une prise de conscience et conduit des établissements à solliciter l'expertise du réseau ARSHS pour l'évaluation et le traitement de leurs archives scientifiques. Des opérations de numérisation ont été montées, des MSH se sont spécialisées dans différents supports. A ce titre les services développés par la MSH René Ginouvès de Nanterre sont souvent cités en exemple : outre un traitement normalisé des archives scientifiques, elle s'attache à numériser des matériaux liés au travail de terrain (carnets, relevés…) en archéologie et ethnologie et collabore avec la MSH d'Aix en Provence à des programmes de traitement des documents sonores et photographiques[8].

Dans le sillage de cette enquête, le MESR a souhaité disposer d'une cartographie des fonds patrimoniaux culturels et scientifiques détenus par les bibliothèques universitaires. Un questionnaire a été conçu dans l'objectif de recenser les collections patrimoniales en attente de signalement, de préciser l'état de conservation de ces collections et d'identifier les actions de valorisation, et plus particulièrement de numérisation. Diffusée en juin 2008, cette enquête est expressément mise en rapport avec celle du réseau des MSH[9]. A partir de ces informations une politique patrimoniale pourrait être bâtie. 109 réponses ont été recueillies, 50 287 manuscrits dénombrés, ainsi que 2 285 fonds d'archives. 32 % des réponses mentionnaient un projet de numérisation[10].

A notre connaissance, cette enquête n'a connu aucune suite. L'état des lieux des opérations de numérisation que le GIS Collex Persée doit mener se situe dans une perspective différente puisque s'attachant à la nature scientifique, et non patrimoniale, du projet, ces deux aspects pouvant néanmoins se recouper par endroits. Il devra recenser les opérations conduites par des bibliothèques, mais aussi celles menées par des équipes de chercheurs, le réseau de MSH étant très actif dans le domaine de la numérisation.

### *2.2.2. Les archives scientifiques nativement numériques*

Cependant, là aussi l'essor du numérique modifie en profondeur la production, la collecte, la diffusion et la conservation des archives scientifiques, désormais nativement numériques, sans que rien de systématique n'ait été entrepris pour organiser la gestion de ces données. Les courriers électroniques mêlent allègrement le privé et le professionnel, comme cela se produisait déjà avec les lettres imprimées. Les fichiers sont enregistrés sur des disques durs ou

[8] L'article de Serge Wolikow publié dans la revue Histoire@Politique à propos de cette enquête et le mémoire de Guillaume Delaunay sur « les archives scientifiques en sciences humaines et sociales : état de l'art » apportent des informations précieuses sur la genèse de la problématique des archives scientifiques en SHS.

[9] Voir BBF 2008, n°6 article de Valérie Néouze

[10] http://www.patrimoineecrit.culture.gouv.fr/files/jpe/2011/MOIlliano.pdf

des clefs USB, rarement sur des serveurs. L'expression « vrac numérique » est parfois employée pour décrire ce désordre. La difficulté nouvelle réside dans l'obsolescence rapide des supports de lecture et des formats de données. Tant que le numérique se bornait au rôle de format de substitution, cette caractéristique n'entraînait pas de catastrophes, elle peut désormais causer la perte irrémédiable de contenus nativement numériques si l'on ne veille pas à leur duplication et à leur conversion régulières. Depuis 2008, les feuilles de route successives des Très Grandes Infrastructures de la Recherche visent à proposer des solutions pour la conservation pérenne des données de la recherche, que ce soit à travers Adonis et Corpus qui ont fusionné pour devenir Huma-Num ou Progedo (Protection et Gestion des Données pour les sciences humaines et sociales). La BSN comportait aussi à sa création un segment sur l'archivage pérenne, englobé par la suite dans l'infrastructure Collex-Persée.

Ces fonds nativement numériques se juxtaposent aux archives sur supports physiques et ne s'y substituent pas. Les archives scientifiques deviennent de plus en plus hybrides et donc complexes.

Que faire de ces archives nativement numériques ? Les réflexions en cours à la **BDIC (Bibliothèque de documentation internationale et contemporaine) et au Campus Condorcet** ouvrent des perspectives particulièrement intéressantes. Les membres fondateurs de ce dernier ont formellement confié au Grand équipement documentaire une mission sur leurs archives scientifiques. Sept d'entre eux (sur 9) lui demandent de gérer toute la chaîne archivistique de la collecte à la communication, en passant bien sûr par la conservation. 4,2 km de documents ont été recensés. La conservatrice responsable du projet « archives » estime indispensable de définir en concertation avec les chercheurs une politique de collecte et des procédures de dépôt pour les laboratoires implantés sur le campus. Cela suppose également des échanges avec les Archives nationales sur les questions de versement.

Depuis ses origines, la BDIC collecte, conserve et met à disposition de ses lecteurs des fonds d'archives privées, qui lui sont donnés ou confiés en dépôt. Cette démarche s'inscrit dans le cadre de sa politique documentaire. Depuis le début des années 1990, elle produit et sauvegarde également des archives audiovisuelles et orales. Ces documents sont signalés dans Calames. Conçus comme des matériaux de la recherche, ils donnent la parole à des témoins et acteurs de l'Histoire. Cette collection est régulièrement enrichie grâce aux travaux des chercheurs, des associations et aux collectes de sources orales que la BDIC met en place avec divers organismes. Le passage au numérique des producteurs qui déposent régulièrement des archives a conduit la Bibliothèque à se pencher sur la question de la gestion des archives nativement numériques. Elle teste avec la Ligue des droits de l'homme (qui lui a déposé ses archives) des procédures de dépôt de façon à déterminer une arborescence, des règles de nommage des fichiers, des principes de description sommaire des lots, des formats d'entrée et des critères de tri. Parallèlement, elle a adhéré au *Consortium Archives du monde contemporain* d'Huma-Num qui comprend un pôle thématique *Archives de la recherche et des chercheurs.* Dans ce périmètre ont lieu des discussions sur la description des archives audiovisuelles en EAD qui devraient aboutir à un code de bonnes pratiques. La BDIC teste des outils développés par Huma-Num pour stocker les fonds à traiter et pour sécuriser leur

conservation. Elle a d'autre part noué un partenariat avec l'Institut national de l'audiovisuel (INA) et procède à des collectes complémentaires sur le web.

### *2.2.3. Les archives d'Internet*

Internet constitue à ce jour pour la recherche un gisement prodigieux. Depuis la loi DAVDSI de 2006 et son décret d'application de 2011, la BNF et l'INA, par extension du champ du dépôt légal, ont pour mission de collecter, conserver et communiquer les sites Internet du domaine français. Elles conjuguent une collecte large entièrement robotisée sur des millions de sites et une collecte ciblée à laquelle contribuent d'autres bibliothèques, en particulier celles habilitées à recevoir le dépôt légal imprimeur en région. Ces bibliothèques sont aussi les seuls lieux hors Ile-de-France où ces archives numériques peuvent être consultées. La Bibliothèque nationale et universitaire de Strasbourg (BNUS) est la seule bibliothèque universitaire à ce jour où les chercheurs peuvent accéder à ces archives par ce qu'elle bénéficie du statut « BDLI ». L'INA cherche à élargir les possibilités de consultation à d'autres établissements de l'ESR.

La base de données de l'INA est alimentée par 14 000 sites web, 20 000 comptes réseaux sociaux et des plateformes vidéo. Elle a été co-construite par des chercheurs et des bibliothécaires/documentalistes au sein d'ateliers du web créés en 2009 afin d'anticiper les usages de la recherche et de réfléchir aux enjeux épistémologiques. Les métadonnées jouent dans cette perspective un rôle majeur. En effet, le web se caractérise avant tout comme un flux continu que l'archivage vient figer dans un état. Cette stabilisation est indispensable à la vérification des sources et à la reproductibilité des démarches scientifiques. Mais elle rompt le lien avec un contexte et une expérience du web que les métadonnées doivent s'attacher à restituer. D'où un énorme travail de description qui doit accompagner la collecte. Des outils de visualisation qui affichent en particulier la distribution temporelle favorisent l'émergence d'hypothèses scientifiques à tester. La BNF a inscrit dans son programme de développement 2016-2019 la préfiguration d'un service de fourniture de corpus numériques qui permettra d'extraire des archives du web, mais aussi des autres collections numériques de la BNF, des ensembles de ressources cohérents, à des fins de fouille de données par des équipes de chercheurs.

# 3. Quels facteurs favorisent la transformation numérique des ressources documentaires en SHS ?

## 3.1. Pourquoi la transformation numérique des ressources documentaires va s'accentuer

### *3.1.1. Le travail du chercheur a changé*

Les humanités, lettres, sciences humaines et sociales, comme les autre sciences, se sont acclimatées à un milieu désormais numérique de bout en bout, même si elles ont suivi leur propre rythme. Elles pratiquent l'e-science à toutes les étapes de leur travail : repérage de sources, de données, recherche et gestion bibliographique, traitement des données, constitution de bases de données, écriture recourant à des cartes heuristiques ou des logiciels d'écriture collaborative, communication scientifique sous forme de blogues, de carnets de recherche, dépôt dans une archive ouverte, présentation à l'aide de divers logiciels …Dans un article de la Revue d'histoire contemporaine, intitulé « Faire de l'histoire à l'ère numérique : retours d'expériences », deux chercheurs de l'Ecole des hautes études en sciences sociales (EHESS) décrivent comment le métier d'historien a été transformé par le numérique qui est à la fois instrument, méthode, terrain et objet de recherche. Hélène Bourdeloie, maître de conférences en sciences de l'information et de la communication, étudie, elle, « ce que le numérique fait aux sciences humaines et sociales ». Les outils mis à disposition par le numérique induisent des méthodes adaptées qui réactivent les débats épistémologiques sur la construction de la science et son statut, sur l'objectivité du chercheur et des données produites. Ces discussions cristallisent autour des humanités numériques qui soulèvent à nouveau frais la question de la place de la technique dans l'élaboration des problématiques de recherche. Les humanités numériques décloisonnent les disciplines en suscitant des échanges sur les technologies utilisées en commun. Elles définissent aussi de nouveaux champs d'étude qui dépassent les frontières classiques des disciplines : « cultural studies », études de genre, études aréales…, mais aussi celles des pays et des langues.

### *3.1.2. La notion de publication évolue*

L'article ou le chapitre au format pdf restent la production scientifique par excellence, mais les projets de recherche génèrent aussi maintenant d'autres produits tels que des bases ou des jeux de données qui ne sont actuellement reconnus qu'en tant que matériaux complémentaires et pourraient jouer à l'avenir un rôle dans l'évaluation de la recherche aussi important que l'article qui énonce les résultats obtenus à partir de ces données si la reproductibilité de ces résultats devient une préoccupation majeure. Ces jeux de données sont publiés soit par l'éditeur s'ils ne sont pas trop volumineux, soit par exposition dans un entrepôt de données tel que Zenodo développé par le CERN pour l'Union européenne. L'écriture et la publication scientifiques se détachent petit à petit du modèle du texte imprimé.

### *3.1.3. Le travail du bibliothécaire a changé*

La prévalence de l'expression « ressources documentaires » sur le terme « collection » employé pour désigner les documents rassemblés dans une bibliothèque révèle un changement complet de paradigme : une ressource implique un but pour lequel elle est mobilisée. Elle est toujours « ressource pour ». Elle ne se suffit pas à elle-même, elle requiert des services qui permettent de l'exploiter. Alors qu'une collection est constituée en fonction d'un idéal et se suffit à elle-même. Le rôle du bibliothécaire s'en trouve transformé : son travail ne se centre plus sur la collection, mais sur l'offre de services qui doit accompagner ces ressources documentaires si le bibliothécaire prend en compte les modalités de travail du public qu'il dessert. La diversification de ces ressources détaillée plus haut étend le périmètre de déploiement de ces services. Il englobe désormais les données de la recherche dans la mesure où celles-ci doivent être enrichies de métadonnées pour être visibles et accessibles.

### *3.1.4. Le rapport de la société à la science a évolué*

Internet offre de nouveaux vecteurs à la communication scientifique, mais aussi à la diffusion des résultats de la recherche et ces possibilités d'y accéder ne laissent pas les citoyens indifférents. Les taux de consultation dont bénéficient certains articles en accès libre de portails tels que Cairn ou Revues.org démontrent cet intérêt. L'utilité sociale des SHS est aujourd'hui réaffirmée à travers des mesures gouvernementales qui visent à faciliter les transferts entre équipes de recherche et acteurs politiques de façon à éclairer l'action publique grâce aux apports de la recherche en SHS. Le rapport « recherches sur les radicalisations » commandé à l'Alliance Athena par le Secrétaire d'Etat à l'Enseignement supérieur et la recherche en novembre 2015 à la suite des attentats perpétrés à Paris et remis le 3 mars 2016, illustre cette volonté de recréer des passerelles entre le monde de la recherche et celui des politiques publiques. L'engouement qui s'exprime pour la science participative relève de la même dynamique de démocratisation de la recherche en sciences humaines et sociales. Celles-ci tendent à opérer une critique sociale, mais aussi à donner à chaque citoyen les moyens d'effectuer la même démarche. Claire Lemercier énonce très clairement la question qui sous-tend l'ouverture de la science : « pour qui écrivons-nous ? », tel est le titre de sa contribution au numéro que la Revue d'histoire moderne et contemporaine a consacré en 2015 à « Economie et politique de l' « accès ouvert » : les revues à l'âge numérique ». Cette question fondamentale permet de trancher le débat qui oppose les tenants des réseaux sociaux académiques, tels qu'Academia ou Research Gate, à ceux des entrepôts d'archives ouvertes institutionnelles ou non. Les premiers nécessitent une identification et ne sont donc accessibles qu'aux seuls chercheurs, les seconds sont ouverts à tous. L'étude du GFII en 2009 notait déjà que « dans l'édition SHS, la question du libre accès est à aborder essentiellement dans la perspective d'un ré-élargissement des audiences de la recherche et d'une restructuration des flux de communication scientifique allant de pair avec un décloisonnement des frontières institutionnelles qui rigidifient aujourd'hui la notion de « communauté de recherche » en SHS . En STM la question de la cherté des tarifs d'abonnements, des marges

bénéficiaires élevées des grands éditeurs internationaux… jouent encore un rôle non négligeable dans le débat sur l'Open Access » (rapport de synthèse page 9). Ce débat ne se pose donc pas de la même façon en STM et en SHS. C'est pourquoi les solutions préconisées diffèrent : le modèle « gold » traite la question des coûts, prépondérante pour les chercheurs en STM, au moyen des APC[11]. Le modèle « green » règle la question de l'accès, prépondérante pour les chercheurs en SHS, grâce aux entrepôts d'archives ouvertes institutionnelles ou pas.

#### *3.1.5. L'édition scientifique change*

En revanche, les difficultés que connaissent les revues en SHS résultent directement du développement du web qui déstructure les maillons de la chaîne éditoriale. L'étude du GFII en distingue cinq (page 13, rapport de synthèse) : l'enregistrement (qui a publié quoi et quand), la validation-certification par un comité de lecture, la mise en forme-enrichissement, la diffusion, la conservation. A l'heure du numérique, seule la troisième étape reste l'apanage d'un éditeur. Il définit une ligne éditoriale, sélectionne des thèmes, crée une communauté de contributeurs autour d'un projet éditorial. Une revue dépourvue de cette valeur ajoutée est simplement publiée. Actuellement, cette valeur ajoutée est rémunérée par un modèle économique que le numérique sape en permettant la redistribution des autres maillons pris en charge par des plateformes et les institutions scientifiques elles-mêmes.

Cette mutation en profondeur du contexte général de la recherche entraîne inéluctablement une transformation de ses rapports à la documentation. Il est attendu que celle-ci s'insère dans l'écosystème numérique qui environne désormais le chercheur en SHS. Même s'il part de sources imprimées, son premier réflexe sera de les scanner ou mieux de les numériser. Si le corpus est vaste, il demandera un financement dans le cadre de son projet pour mener à bien cette opération.

### 3.2. Comment favoriser la transformation numérique des ressources documentaires en SHS ?

#### *3.2.1. Développer une offre de services autour des ressources documentaires à partir des besoins des chercheurs*

L'étendue, la profondeur et la multiplicité des gisements documentaires existants rendent utopique une transposition numérique globale. Il paraît beaucoup plus réaliste de se fonder sur les besoins des chercheurs. Cela suppose une coopération étroite entre bibliothécaires et chercheurs. Le projet Fonte Gaïa peut à ce titre servir d'exemple : chercheurs et bibliothécaires de l'Université de Grenoble Alpes ont conjugué leurs efforts afin de créer une communauté d'acteurs souhaitant « s'investir dans un réseau d'échange d'information et de

[11] Articles processing charges.

débat d'idées autour de l'italianisme contemporain et international » à l'heure des humanités numériques. Ces collaborations se concrétisent sous la forme d'un blogue et d'une bibliothèque numérique de référence constituée par moissonnage de contenus dispersés sur différents sites, sélection et numérisation de nouveaux corpus sous la responsabilité d'un comité scientifique, établissement de commentaires qui viennent enrichir les éditions numériques. « Une plateforme de diffusion est actuellement en construction à l'université de Grenoble. L'ensemble de ses fonctionnalités seront développées selon les préconisations d'un panel d'utilisateurs. ».

L'université de Lille 3 est aussi engagée dans des projets d'humanités numériques qui reposent sur des collaborations entre chercheurs, bibliothécaires et ingénieurs de recherche : le projet TALIE (Traditions de l'Antiquité à Lille et dans l'Eurorégion) est dédié à la valorisation d'œuvres de l'Antiquité gréco-romaine présentes dans les fonds patrimoniaux des bibliothèques de la région Nord-Pas-de-Calais-Picardie, sous la forme des manuscrits ou d'éditions anciennes. Le projet vise entre autres à établir, pour quelques-uns de ces ouvrages, une édition numérique enrichie d'hypertextes et une traduction en français quand l'ouvrage est en langue ancienne. La Maison européenne des sciences de l'homme finance les opérations, l'UMR Halma coordonne et le Service commun de la documentation apporte des collections patrimoniales, des compétences en méthodologie, en bibliographie, et gère les relations avec les autres bibliothèques-ressources du territoire.

La refonte en cours des CADIST (Centres d'acquisition et de diffusion de l'information scientifique et technique) s'inscrit précisément dans cette perspective. L'infrastructure de recherche Collex-Persée qui se substitue à l'ancien dispositif part des attentes des chercheurs pour repenser l'offre de services autour de la documentation de niveau recherche. L'enquête menée par le Bureau van Dijk dans le cadre de cette réflexion a souligné l'existence de demandes spécifiques en fonction des disciplines. Les chercheurs en SHS expriment à 67 % le besoin de disposer de davantage de documents textuels numériques, à 54 % de documents textuels imprimés, à 24 % de données et à 19 % d'archives de chercheurs. Ils classent en services devant être prioritairement proposés : la numérisation à la demande, la veille documentaire avec alerte, l'achat d'articles hors abonnement à la demande (« pay-per-view »). **Les résultats de cette enquête doivent nourrir une stratégie de numérisation claire, non pas tant sur les corpus à numériser** parce que ce sont les chercheurs et les instances scientifiques qui sont les mieux placés pour déterminer en fonction de leurs axes de recherche ce qui doit l'être, **mais sur le signalement, l'enrichissement des données, la diffusion, l'archivage de ces corpus et le statut juridique des documents numérisés.** Persée a entamé une réflexion à ce sujet qui doit être élargie**. Les rôles des différents protagonistes existants doivent être clarifiés et articulés.** L'infrastructure Collex-Persée peut lancer des appels à projets de numérisation en lien avec des programmes de recherche et les financer. **Le plus important est qu'elle impose un signalement dans Calames ou le Sudoc en fonction des types de documents, une normalisation de l'encodage des données, des métadonnées et des identifiants ainsi qu'une diffusion sur la plateforme Persée pour certains types de documents et un moissonnage par Isidore**. Elle doit en outre rendre accessible à tous les corpus remplissant les critères précédemment indiqués une **solution**

**d'archivage pérenne à coût modéré voire gratuit pour la communauté de l'ESR**. Pour les projets qui seraient conduits en dehors du cadre Collex-Persée, il convient d'ériger ces principes au moins en bonnes pratiques, comme l'a fait Huma-Num dans son Guide des bonnes pratiques. La BNF est membre du GIS Collex-Persée en tant qu'opérateur national, il paraît donc naturel d'articuler Persée et Gallica afin d'accroître la visibilité des corpus mis en ligne.

Les documents n'appartenant pas encore au domaine public soulèvent des questions juridiques qui nécessitent souvent la signature d'une convention avec les ayants droit. Dans le cadre d'une politique nationale, il conviendra de **lever les obstacles juridiques qui entravent la fourniture de documents sous forme numérique**, que les documents soient numérisés ou nativement numériques.

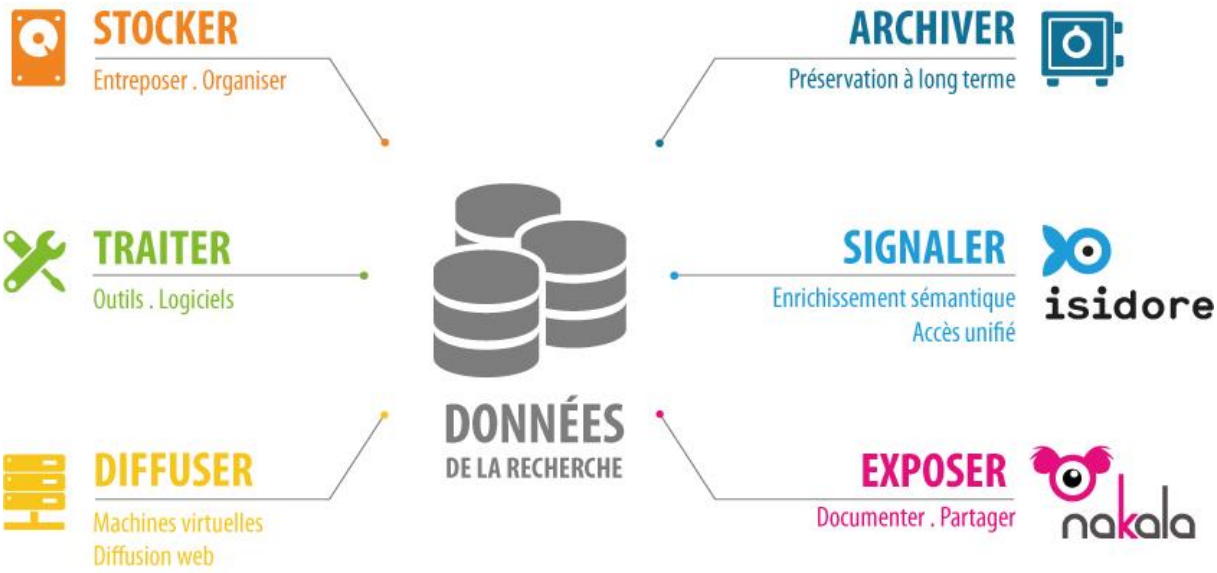
**La numérisation à la demande** réclamée par les chercheurs peut se pratiquer en interne si la bibliothèque qui détient le document convoité possède aussi les équipements nécessaires. La BCU de Clermont, le SICD de l'Université de Grenoble Alpes et la BDIC offrent ce service à leurs usagers pour les documents libres de droit, moyennant paiement. La demande est déposée grâce à un formulaire en ligne. La BIU Santé, la BNUS et très récemment le réseau des Bibliothèques Universitaires de Toulouse-Midi-Pyrénées ont opté pour une sous-traitance partielle déléguée à EOD (Ebooks on demand). Les documents qui peuvent faire l'objet d'une numérisation sont identifiés dans le catalogue de la bibliothèque par un pictogramme. Après un clic, l'usager est invité à créer un compte sur la plateforme EOD à partir duquel il pourra formuler sa demande, suivre sa commande et payer en ligne après acceptation du devis qui lui sera soumis. La numérisation s'effectue toujours à la bibliothèque. Un lien est envoyé à l'usager pour télécharger le fichier. Les documents numérisés à la demande peuvent ensuite être mis à la disposition de tous. Ce type de prestation trouve facilement sa place dans une politique de site qui vise à mutualiser des services aux chercheurs. Le matériel et l'acheminement des originaux ne soulèvent pas de problèmes insurmontables. Il faut veiller à tenir des délais de fourniture aussi rapides que possible. Seules les questions juridiques afférentes aux documents hors domaine public peuvent entraver une initiative de cette nature.

Le développement des services aux chercheurs, en phase avec leurs attentes, nécessite une **ouverture des équipes de recherche** qui permettent à des bibliothécaires d'y participer dès la naissance d'un projet. Les questions déterminant la structuration des données et des métadonnées doivent être posées dès ce stade : les objectifs du programme de recherche induisent des modalités d'exploitation des données qui elles-mêmes impliquent des décisions en matière de structuration des données et des métadonnées. On évite ainsi de se voir imposer des choix par défaut : que peut-on faire avec les données dont on dispose et les conditions d'utilisation fixées par le propriétaire ? Le dialogue entre chercheurs et bibliothécaires se noue aussi autour des enjeux des référentiels retenus. Il faut ensuite préciser si les données, que ce soient des documents numérisés, des bases de données ou des jeux de données, seront diffusées et de quelle manière. Leur référencement doit être optimisé. Leur durée de vie dépend, elle, des choix arrêtés pour leur sauvegarde. Un archivage pérenne requiert une conformité à des standards internationaux tels OAIS qui doivent être pris en compte très tôt

dans la gestion du projet. Autant de sujets qui gagnent à être débattus au sein d'une équipe alliant autour de chercheurs des compétences variées : bibliothéconomiques, archivistiques, informatiques, juridiques…

### *3.2.2. Les services autour des données de la recherche*

La démarche suivie par Huma-Num pour bâtir son offre de services fournit un modèle aux bibliothèques. Cette offre est conçue à partir de la notion de « programme de recherche » et se décline en six actions qui en couvrent le cycle de vie :




Source : site Huma-Num

Ce dispositif technologique s'appuie sur des communautés scientifiques organisées en consortiums qui « réunissent plusieurs unités et équipes de recherche autour de thématiques et d'objets communs pour lesquels ils définissent des procédures et standards numériques partagés (méthodes, outils, partages d'expériences). Ils formulent un programme d'actions pluriannuel pour lequel ils s'engagent à apporter une partie des moyens nécessaires (en particulier en personnel), et pour lequel ils demandent, en outre, un soutien de la part de la TGIR. Ces actions donnent lieu à des livrables. Dans ce cadre, les consortiums sont labellisés et évalués par le comité de pilotage et le conseil scientifique de la TGIR. »

Les consortiums actuellement labellisés sont au nombre de dix :

- CAHIER - Corpus d'Auteurs pour les Humanités : Informatisation, Édition, Recherche
- Archives des ethnologues
- Musica
- CORLI - Corpus, Langues et Interactions
- Consortium 3D-SHS
- MASA - Mémoires des archéologues et des sites archéologiques

- COSME - Sources Médiévales
- IMAGEO - Cartes et photographies pour les géographes
- ArchiPolis - Archives des sciences sociales du politique
- ArcMC - Archives des mondes contemporains

La TGIR Huma-Num consacre 60 % de son budget à ces consortiums qui regroupent, outre des chercheurs, des documentalistes, des archivistes et quelques bibliothécaires. Des chercheurs étrangers intègrent ces consortiums par le biais du réseau européen DARIAH. Ces consortiums fédèrent environ 300 personnes. Les Maisons des sciences de l'homme jouent un rôle important dans l'animation de ces communautés. Elles organisent en particulier des formations et des ateliers.

Cependant, des regroupements d'établissements, dans le cadre d'une politique de site, cherchent à se doter, au moins partiellement, d'une offre de services semblable qui s'articulerait autour de ressources documentaires locales. Ne serait-il pas plus intéressant **d'élargir les communautés auxquelles Huma-Num ouvre ses services**, ce qui suppose d'allouer des moyens supplémentaires à cette infrastructure, et de créer des plateformes locales de valorisation des ressources documentaires moissonnées par Isidore ? Des bibliothèques universitaires, des services IST de COMUE pourraient ainsi recourir aux outils développés par Huma-Num au profit certes de leur communauté locale, mais avec une visibilité nationale et internationale.

La fouille de textes autorisée par la loi République numérique promulguée le 7 octobre 2016 illustre à merveille la nécessité de cette synergie de compétences et la place particulière qu'y tiennent celles des bibliothécaires. Cette fouille s'applique à des corpus documentaires dont il faut d'abord vérifier qu'ils rentrent dans le cadre légal. Ils doivent ensuite être délimités, normalisés et harmonisés de façon à pouvoir être manipulés par les outils de fouille. Les bibliothécaires y contribuent par leur expertise en matière de gestion documentaire et jouent un rôle de médiateur qui leur est familier en orientant les chercheurs vers des logiciels et des formations. Le projet ISTEX montre cette synergie à l'œuvre : les bibliothécaires travaillant au sein du consortium Couperin participent activement à la sélection de ressources documentaires numériques (archives de revues, livres électroniques, corpus de textes). Ils recueillent et analysent les besoins des chercheurs, étudient les offres des éditeurs et évaluent les ressources. Celles-ci sont acquises grâce à des financements « Initiative d'excellence » du programme « Investissements d'avenir » du ministère de l'Enseignement supérieur et de la Recherche. Ces ressources documentaires sont ensuite déposées sur une plateforme nationale où elles sont exploitées par un ensemble de services utilisant notamment des technologies de fouille de contenus. Le dernier appel à propositions lancé par ISTEX cible directement des ressources documentaires en langue française peu présentes jusqu'ici, ce qui devrait permettre aux Sciences humaines et sociales de renforcer leurs corpus sur cette plateforme et d'utiliser ensuite les services de fouille qui permettent de « libérer des connaissances cachées dans les textes » et ouvrent donc de nouveaux horizons à la recherche.

### *3.2.3. Les services autour des publications*

Le passage au numérique implique le développement de services autour des ressources documentaires. Elles ne peuvent se suffire à elles-mêmes dans ce contexte. Elles nécessitent des outils non seulement pour les lire (ordinateurs, tablettes, liseuses, téléphone « intelligent »…), mais aussi pour les utiliser. Cela vaut quand elles forment corpus, mais aussi quand elles sont envisagées à l'unité, en tant que « publication ». Les plateformes sur lesquelles ces publications sont diffusées proposent des services aux éditeurs, aux bibliothèques et aux utilisateurs. Le tableau ci-dessous permet de comparer l'offre des trois plateformes françaises les plus importantes en SHS.

| | **CAIRN** | **OpenEdition** | **PERSEE** |
|---|---|---|---|
| **Services aux revues/éditeurs :** Numérisation ? chaîne de publication électronique ? | Numérise ou convertit les fichiers ayant servi à la fabrication papier en XML (DTD Erudit) et format « page » (PDF, SWF) | Chaîne de publication électronique mise à disposition sous Lodel et au format XML TEI. Pour les livres, programme de soutien à la numérisation | Numérise (XML TEI pour le contenu) |
| Hébergement ?<br>diffusion ?<br>distribution ?<br>commercialisation ? | oui<br>oui<br>oui<br>oui | oui<br>oui<br>oui<br>oui | oui<br>oui<br>non<br>non |
| Archivage ? | Architecture répliquée, Discussion avec BNF actionnaire de Cairn | Système de sauvegarde du CCSD.Projet en cours avec le CINES via Huma-Num | Archivage pérenne au CINES |
| **Services aux bibliothèques :** Unité documentaire ? | Article/chapitre>numéro/ ouvrage>revue/collection | Idem + annonce (Calenda), billet (Hypotheses) | Idem +communication pour actes de colloque, accent mis sur la collection comme ensemble structuré de documents |
| Entrepôt de métadonnées pour moissonnage ? Format d'exposition des métadonnées ? | Moissonnage OAI-PMH possible. Métadonnées exposées aux formats Dublin Core et METS. Sets : bouquets commerciaux, revues/collections | Moissonnage OAI-PMH possible. Métadonnées exposées aux formats : Dublin Core simple et qualifié, METS, TEI. Sets : revues, éditeurs de livres, carnets | Moissonnage OAI-PMH possible. Métadonnées exposées en Dublin Core, MODS, marcXML, METS et MAD pour les autorités Ouverture d'un triplestore en février 2017 (RDF) |

| | **CAIRN** | **OpenEdition** | **PERSEE** |
|---|---|---|---|
| TEI est-il utilisé pour la description du texte intégral ? | Non, mais a établi des programmes de conversion entre XML Erudit et TEI et inversement | Oui, pour les articles et les chapitres | Oui, à l'avenir utilisera EAD pour documents d'archives |
| récupérer les notices bibliographiques ? quel format ? | Oui, Unimarc et Marc 21 | Oui, Unimarc et Marc 21 | Même réponse que pour moissonnage |
| Partenariats pour référencement dans des catalogues collectifs, des bases de connaissance, des moteurs de recherche ? lesquels ? | Google Scholar, Isidore<br>Ebsco, ProQuest, ExLibris, BACON<br>SUDOC, Worldcat<br>Mir@bel, Repères(Canada), MIAR (Espagne)<br>Scopus, Wos, Wikipédia<br>PMB, Archimed… | Google Scholar, Isidore<br>Ebsco, ProQuest, ExLibris, Elsevier (science direct)<br>SUDOC, Worldcat | Google, Bing, Google Scholar (implémentation de schema.org)<br>Ebsco, ProQuest, ExLibris<br>Sudoc |
| Utilisation du format Kbart pour transfert de listes aux partenaires ? | Fichiers Kbart créés en collaboration avec Abes | Oui, pour les revues et les livres | Fichiers Kbart créés en collaboration avec Abes |
| Les métadonnées sont-elles alignées sur des référentiels ? lesquels ? | Réflexion en cours sur IdRef | Bisac (subject headings book industry study group) | IdRef, GBIF (sciences de la vie et de la terre), geonames |
| Quelles licences d'utilisation accompagnent les services et les documents? | Numérisation, traitement de fichiers payant pour les éditeurs. Accès en ligne au texte intégral payant avec barrière mobile de 2 à 5 ans.<br>Les services sont gratuits (alertes, notices…)<br>Les éditeurs choisissent régime droit d'auteur ou licence Creative commons | De plus en plus de revues adoptent des licences Creative commons .<br>Les services proposés dans le cadre de l'offre « freemium » sont payants. | Tous les services aux revues, aux bibliothèques et aux internautes sont gratuits.<br>Une licence CC BY NA NC est associée aux documents.<br>Réflexion en cours sur Public Domain Mark<br>Triplestore : licence CC BY NC SA pour jeux de données.<br>Réflexion en cours sur Etalab |
| **Services aux utilisateurs :**<br>Quels types de documents sont proposés ? avec quelle granularité ? | 450 revues, 6 500 ouvrages de recherche, 2 100 encyclopédies de poche, dizaine de titres de presse spécialisée. | 451 revues, 3 510 livres, 1878 carnets de recherche, plus de 34 000 événements signalés dans Calenda | 645 567 documents diffusés (revues, livres, actes de colloque et sources primaires (archives, cartes, photos…) |

| | **CAIRN** | **OpenEdition** | **PERSEE** |
|---|---|---|---|
| Quels champs disciplinaires sont couverts ? | Sociologie<br>Sciences politiques<br>Economie-Gestion<br>Droit<br>Info-Com<br>Histoire<br>Géographie<br>Arts<br>Lettres et linguistique<br>Sciences de l'éducation<br>Psychologie<br>Santé publique<br>Travail social | SHS, énergie, environnement, santé,<br>Collège de France et IRD (disciplines sciences exactes aussi) | anthropologie ; archéologie&art ; démographie ;droit; écologie;économie-gestion;études classiques aires culturelles ; études régionales; géographie ; histoire ; épistémologie ; histoire des sciences et techniques; linguistique ; littérature; philosophie ; psychologie ; religion et théologie ; sciences de l'éducation ; science politique ; science de l'information et de la communication ; sciences et vie de la terre ; sciences et techniques, sociologie |
| Possibilités d'alertes ou d'abonnement à des services ? | Alertes sur nouveautés, flux RSS par revue, discipline, type de contenus | Alertes sur les nouveautés sur chaque plateforme | Non, services supprimés parce qu'inutilisés.<br>réseaux sociaux |
| Quels formats de données ? quotas ? DRM ? | PDF, pas de DRM | PDF et Epub pour revues et livres, pas de DRM | PDF enrichi, pas de DRM |
| Export de notices ? quels formats ? | Zotero, Endnote, Refworks | Zotero | BibTex, Refworks, RIS, compatible Zotero |
| Conserver un historique de ses recherches ? | Oui, avec ou sans création de compte personnel | Non, pas de fonctionnalité développée | Non, a été supprimé, pas utilisé |
| Existe-t-il une version anglaise de l'interface ? autres langues ? | Interface spécifique pensée pour les non francophones (revues).<br>Amélioration des métadonnées en anglais produites par les revues.<br>Moteur de recherche dédié au fonds anglais (4500 articles). | Moteur de recherche général : interface en anglais<br>Moteur OpenEdition Books : anglais, espagnol, italien, allemand<br>Moteur Revues.org : anglais | Non, prévu pour 2017-2018 : anglais, allemand, espagnol |
| Quelle proportion de documents en français ? | 97 %<br>Sur Cairn-int.info : uniquement en anglais (traductions) | Toutes plateformes confondues : 69 %<br>Livres : 2762 en français, 301 en | 90% |

| | | | |
|---|---|---|---|
| | | espagnol, 238 en anglais, 79 en italien, 59 en allemand | |
| | **CAIRN** | **OpenEdition** | **PERSEE** |
| Recherche en texte intégral ? | oui | oui | oui |
| Comment les usagers sont-ils tenus informés des évolutions ? quelle prise ont-ils sur ces évolutions ? | Sites d'actualité, présence sur Twitter, Facebook, Scoop.it<br>Focus group utilisateurs une fois par an | Newsletter, mailing, blog du Cléo, réseaux sociaux<br>Comité d'utilisateurs, assemblée des revues, assemblée des carnetiers, consortium des éditeurs | Présence sur Twitter et Facebook.<br>Prévu en 2017 : site institutionnel, comité d'utilisateurs (chercheurs, éditeurs, professionnels IST) |
| Quelles statistiques et mesures d'usage ? | Fichiers format COUNTER pour les institutions abonnées | Fichiers format COUNTER V4 et SUSHI pour les institutions abonnées à l'offre freemium.<br>Statistiques de fréquentation consultables par tous (AWStats) | Fichiers format COUNTER pour les éditeurs et pour les institutions qui en font la demande.<br>Statistiques de fréquentation consultables par tous (AWStats) |
| DOI attribué ? intégré à la notice ? | Dépose un DOI auprès de Crossref pour ceux qui n'en ont pas.<br>Est intégré aux exports | DOI attribués et intégrés aux notices | DOI attribués et intégrés aux notices |
| Quels services prévus en 2017 ? | Refonte ergonomique<br>Ajouts de rapports COUNTER (JR2, BR3)<br>Nouveau flux d'envoi de métadonnées aux éditeurs de logiciel<br>Intégration de la technologie EduGain | Transformation de Revues.org en OpenEdition Journals<br>Robot d'annotation de bibliographies | Ouverture du triplestore data.persee.fr<br>Traduction de l'interface<br>Site institutionnel<br>Comité d'utilisateurs |
| Quels services à l'étude ? | Référentiel IdRef<br>Conformité COUNTER V4 et SUSHI<br>Consultation sur terminaux mobiles | OpenEdition Lab mène des recherches notamment sur la base du TDM | Mise à disposition de données pour de la fouille de texte |

Les choix techniques opérés par les trois plateformes sont très proches. Les modèles économiques diffèrent en revanche sensiblement. Persée est entièrement gratuit parce que

subventionné par des fonds publics, Cairn suit une logique commerciale, OpenEdition allie des financements publics à la facturation de services et cherche à bâtir un modèle économique viable. La volonté de s'internationaliser en s'ouvrant à une consultation et même à une publication non francophone se dessine à travers le développement d'une interface en anglais. Elle s'affirme très nettement chez OpenEdition qui diffuse déjà 31 % de ses livres dans une autre langue que le français. La stratégie de Cairn en la matière consiste à faciliter la découverte d'articles français par des non francophones en améliorant les métadonnées en anglais et en proposant une interface conçue dans cette perspective. OpenEdition et Persée tendent aussi à sortir du champ des SHS stricto sensu pour accueillir des publications relevant des sciences de la vie et de la terre, voire des STM. Le plus remarquable est la diversification des documents hébergés : ces trois plateformes ont commencé par les revues et s'y sont tenues pendant plusieurs années. Puis les encyclopédies de poche et les ouvrages de recherche ont fait leur entrée sur Cairn. Les livres, carnets de recherche et « fiches » événements se sont agrégés à Revues.org chacun sur leur plateforme articulée au moteur de recherche OpenEdition. Persée a élargi son spectre aux différents types de publications scientifiques et même aux sources primaires (archives, cartes, …). Une logique de service préside visiblement à ces extensions.

### *3.2.4. La politique documentaire dans un contexte d'économie de services*

Les grandes maisons d'édition internationales (Elsevier, Springer, Informa…) et les distributeurs-agrégateurs fondent désormais leur modèle économique davantage sur des services articulés autour d'une plateforme que sur les publications elles-mêmes. Certains vont jusqu'à donner libre accès à ces dernières et tentent de trouver un équilibre économique par la vente de prestations. C'est le cas d'OpenEdition avec son modèle freemium dont le succès s'étend au-delà des frontières francophones, mais aussi d'Open Book Publishers, d'Open Humanities Press ou d'OAPEN.

Comment élaborer une politique documentaire dans ce contexte ? Le changement de paradigme évoqué plus haut revêt ici une forme très concrète : le bibliothécaire ne constitue plus une collection dont il choisirait les éléments un à un. Le passage de nombre d'éditeurs de revues numériques à une offre documentaire en « bouquets » qui empêchait la sélection à l'unité avait déjà mis à mal ce principe. Dans cette nouvelle configuration, le bibliothécaire choisit une offre de services vers laquelle il oriente son public. Adhérer à une structure de type OpenEdition devient un acte de politique documentaire. Plus de 132 bibliothèques et institutions sont abonnées au programme OpenEdition Freemium selon une grille tarifaire qui prend en compte le PIB par habitant du pays et le nombre d'étudiants en SHS. Cette contribution financière leur ouvre des services tels que l'accès aux fichiers aux formats PDF et epub, à un espace de gestion des acquisitions et des collections, à des suggestions d'acquisition, à des statistiques COUNTER, à un entrepôt OAI-PMH, la fourniture de notices et de métadonnées, des formations, une personnalisation de la barre de navigation et des versions epub et PDF avec le logo de l'institution…

Les questions de politique documentaire sont subtiles parce que certes régies par des principes fortement théorisés en France, mais aussi en prise sur une réalité locale qui l'emporte régulièrement dans les décisions finales. La British Academy for the Humanities and Social Sciences a mené une enquête dont les résultats montrent ce raffinement. Elle a demandé aux bibliothèques de déterminer le degré d'influence de 6 facteurs différents sur leur décision de supprimer un abonnement à une revue. **Le facteur jugé le moins important est la durée de l'embargo** (noté 2 sur une échelle de 1 à 5 où 5 correspond au facteur le plus important). Le critère qui a obtenu l'évaluation la plus haute est la contrainte financière (4,8), viennent immédiatement derrière les besoins des enseignants-chercheurs et un changement dans les cours dispensés par l'institution à laquelle la bibliothèque appartient (4,5). Ce classement souligne l'ancrage dans une communauté d'usagers dont les attentes sont prépondérantes par rapport à une hypothèse qui voudrait que l'on résilie un abonnement quand la durée d'embargo tombe sous un certain seuil. Une offre de services se construit en fonction d'un public.

### *3.2.5. Les services autour de l'édition*

L'apparition de plateformes de diffusion numérique a profondément bouleversé l'offre documentaire. La nécessaire conjugaison des dimensions « services » et « ressources documentaires » a été soulignée à plusieurs reprises. Cette mise en relation qu'opère la plateforme en tant qu'intermédiaire peut s'élever au degré de sophistication d'un écosystème si elle se déploie sous la forme d'un réseau d'interrelations entre services. OpenEdition et un certain nombre d'éditeurs travaillant avec Cairn ont nettement optimisé l'articulation des processus d'édition et de diffusion en adoptant les Méthodes et outils pour l'édition structurée (Métopes) élaborés par la Maison de la recherche en sciences humaines de Caen. Cette évolution suppose cependant de renoncer à fabriquer des livres ou des revues pour produire un contenu éditorialisé qui se présente sous la forme d'un fichier XML TEI encodé en Unicode. Ce fichier devient le pivot de processus visant à décliner ce contenu en différents produits ou services, tel qu'un livre ou un numéro de revue imprimé ou numérique. Le balisage en TEI et la sémantisation à l'aide de référentiels garantissent non seulement une grande fluidité entre la création et la diffusion de la publication scientifique, mais aussi un repérage beaucoup plus aisé pour la citation et l'indexation. L'infrastructure de recherche Numédif qui promeut Métopes a pour ambition d'accroître la visibilité de l'édition scientifique publique française à l'international. Une cinquantaine de structures éditoriales publiques ont été formées par ses soins à ces méthodes et outils.

A l'échelon européen, la création conjointe d'OpenAire qui moissonne des archives ouvertes de niveau institutionnel, disciplinaire ou national et de Zenodo, entrepôt de données de la recherche, afin d'accompagner l'obligation de dépôt en libre accès dans le cadre d'Horizon 2020[12], esquisse un écosystème européen de la publication scientifique. Dans le domaine des sciences humaines et sociales, OpenEdition s'appuie sur son label national d'infrastructure de recherche pour se lancer dans la structuration d'un réseau de partenaires européens autour

[12] Le programme Horizon 2020 de l'Union européenne comporte l'obligation d'assurer le libre accès aux publications issues des recherches qu'il aura contribuées à financer, sous peine de sanctions financières.

d’un projet visant à construire une infrastructure européenne pour les publications de sciences humaines en accès ouvert (OPERAS).

Cette intégration des chaînes de production et de diffusion explique sans doute le développement d’une activité de publication au sein des bibliothèques. L’édition 2016 du Library Publishing Directory recense 115 bibliothèques se livrant à cette activité. Aucune bibliothèque universitaire française ne figure dans cet inventaire. Plusieurs d’entre elles ont néanmoins déjà communiqué sur un service de cette nature, dont la bibliothèque de Clermont Université et celle de l’Université de Bordeaux. Toutes deux ont installé et testé le logiciel libre OJS (Open Journal Systems) qui permet de coordonner les différentes phases du traitement d’un article (soumission, relecture, validation, publication) dans un espace propre à chaque revue. Elles assistent ensuite les chercheurs dans la prise en main et l’utilisation. Des bibliothèques aident aussi les équipes de revues qui souhaitent basculer sur Revues.org, voire participent au balisage XML des articles. Elles sont sollicitées parce que disposant de compétences techniques en matière numérique, mais aussi juridiques et économiques à force de décrypter des licences régissant l’usage de ressources numériques et les modèles économiques de leurs fournisseurs. Cependant, elles savent très bien où s’arrête leur métier et où commence celui de l’éditeur ou de la secrétaire de rédaction. Là encore ce sont des équipes mixtes qui aboutissent au meilleur résultat en alliant leurs différentes compétences.

# Conclusion

Les dépenses documentaires en sciences humaines et sociales, dont le montant était estimé en 2007 à 37,2 millions d'euros selon une étude menée par le GFII, s'élevaient en 2015 à 37,7 millions. L'analyse de la répartition de ce montant par type de documents montre que les ressources documentaires en SHS ont résorbé ou presque l'écart qui les distinguait des autres champs disciplinaires concernant la part du numérique : en SHS, la part des dépenses documentaires dévolue aux **ressources électroniques** est passée de 22 % en 2007 à **42,21 % en 2015,** alors que, toutes disciplines confondues, les ressources électroniques représentent 50 % des dépenses en 2015. L'enveloppe financière allouée aux SHS ayant peu évolué, cet accroissement s'est opéré par déplacement des dépenses **au détriment des revues imprimées qui perdent près de 17 %,** alors que les ouvrages imprimés n'accusent une baisse que de 4,5 % dans le secteur des SHS. Le nombre de titres de revues françaises auxquelles les bibliothèques des universités à dominante SHS sont abonnées sous forme imprimée a diminué de 33,1 % entre 2007 et 2015, le nombre de titres étrangers de 41,6 %. Les titres français ne représentent plus que 31,2 % de la dépense pour les abonnements à des revues imprimées pour ces mêmes bibliothèques, contre 38 % en 2007.

Tandis que le livre imprimé drainait 5,8 millions d'euros, les livres numériques ont bénéficié en 2015 d'une dépense de 639 000 euros dans les bibliothèques des universités à dominante SHS. Autant dire qu'ils sont loin de se substituer à l'imprimé. Cependant, deux phénomènes sont apparus depuis 2009 qui soulignent le rôle prépondérant que la monographie continue à jouer dans la publication scientifique en SHS : les plateformes qui diffusaient initialement des revues en ligne se sont toutes ouvertes aux livres (CAIRN, Persée, OpenEdition, JSTOR…) et la publication en libre accès s'est étendue aux livres, selon des modèles économiques innovants.

La volonté de dresser un état des lieux de la transformation numérique des ressources documentaires en SHS se heurte à un manque de lisibilité de l'offre éditoriale aussi bien que des ressources numérisées. Les revues françaises ont muté vers le numérique puisque sur les 936 revues françaises labellisées par le HCERES seules 13 sont totalement absentes du web. Mais cette évolution ne signifie pas un renoncement à la version imprimée et n'a pas entraîné une concentration massive de l'offre. Deux services, la bibliothèque numérique Gallica et le moteur de recherche Isidore donnent une bonne visibilité aux corpus de l'ESR qu'ils hébergent ou moissonnent. Toutefois, de nombreux corpus numérisés leur échappent. Le GIS Collex-Persée devra clarifier la mission des différents acteurs concernés et porter une politique concertée en matière de numérisation.

Les ressources documentaires utilisées par les chercheurs se sont diversifiées et incluent désormais des archives scientifiques sur support physique, mais aussi nativement numériques. Ces nouveaux gisements documentaires soulèvent de nouvelles questions techniques, juridiques et méthodologiques que des établissements précurseurs s'attachent à défricher. Les données de la recherche font partie de ces archives scientifiques.

La transformation numérique des ressources documentaires ne peut que s'accentuer encore parce que le chercheur travaille maintenant dans un environnement très largement numérique. L'écriture, la publication et l'édition muent, le travail du bibliothécaire suit nécessairement la même voie. L'expression « ressource documentaire » manifeste un changement de paradigme : une ressource ne se suffit pas à elle-même, elle implique une finalité, un service qui permet de l'exploiter. Cette offre de services doit être élaborée à partir des besoins des chercheurs et requiert une ouverture des équipes de recherche à d'autres métiers qui apportent des compétences diversifiées : bibliothéconomiques, archivistiques, informatiques, juridiques…Ces services doivent se déployer autour des données de la recherche, des archives scientifiques, des publications et des processus éditoriaux. Dans ce nouveau contexte d'économie de services, la notion de « politique documentaire » doit être revisitée en profondeur.

# Préconisations

Améliorer la visibilité et la lisibilité de l'offre documentaire numérique, qu'elle soit commerciale ou en libre accès :

- Inciter les éditeurs français en SHS à renforcer la visibilité de leurs revues et ouvrages numériques en concentrant leur offre sur les deux ou trois plateformes qui diffusent ce type de documents en appliquant des normes et standards internationaux depuis le processus éditorial jusqu'au référencement sur Internet, y compris l'archivage pérenne, et en développant le libre accès.
- Dans le cadre du GIS Collex-Persée, dresser un état des lieux des fonds déjà numérisés, clarifier le rôle dévolu aux plateformes Persée, Huma-Num, Gallica, et au catalogue Calames, définir une politique nationale de numérisation, qui établisse des recommandations contraignantes en matière de statut juridique des documents numérisés, de signalement (Calames, SUDOC), d'enrichissement des données (TEI…), d'encodage des métadonnées, d'utilisation de référentiels nationaux ou internationaux, de diffusion (Persée…), de référencement (balisage sémantique, techniques SEO) et d'archivage pérenne. Analyser les obstacles juridiques à la fourniture de documents sous forme numérique et proposer des solutions.

Mutualiser les dispositifs et les services au sein de la communauté de l'enseignement supérieur et de la recherche :

- Engager des discussions entre le MESR et la TGIR Human-Num afin que des équipes scientifiques autres que les dix consortiums existants puissent utiliser les services conçus par Huma-Num, en particulier dans le cas de COMUE qui souhaitent exploiter et valoriser des gisements documentaires locaux.
- Inciter les bibliothèques universitaires à mieux faire connaitre le moteur de recherche Isidore aux étudiants et aux chercheurs de leur institution et à demander le moissonnage de leurs données.
- Développer la numérisation à la demande dans le cadre des politiques de site

Développer les compétences :

- Favoriser un partage d'expérimentations et de pratiques entre bibliothèques qui se lancent dans le développement de services autour des archives scientifiques, des données de la recherche et des humanités numériques en collaboration avec des équipes de recherche.

- Encourager l’ouverture des équipes de recherche à d’autres métiers qui leur apportent des compétences indispensables à tout projet numérique : bibliothéconomiques, archivistiques, informatiques, juridiques …
- S’appuyer sur cette mixité des équipes de recherche pour développer les compétences sur la gestion des données de la recherche, des archives scientifiques, des humanités numériques, et sur la communication scientifique.

Inciter les bibliothèques universitaires à repenser leur politique documentaire en tirant toutes les conséquences des nouveaux modèles économiques qui déplacent la valeur du contenu vers le service.

# Annexes

---

Annexe 1 : Lettre de mission de l'Inspection générale des bibliothèques pour 2016-2017

Annexe 2 : Liste des personnes rencontrées

# Annexe 1

Liberté • Égalité • Fraternité
RÉPUBLIQUE FRANÇAISE

*La Ministre de l'Education nationale, de l'Enseignement supérieur et de la Recherche*

*La Ministre de la Culture et de la Communication*

Monsieur le Doyen,

Les bibliothèques sont sur l'ensemble du territoire d'irremplaçables lieux de lecture pour tous les publics et à tous les âges, de formation et de recherche, d'accès le plus large à la culture et d'éducation artistique et culturelle. A ce titre, elles sont déterminantes pour la vitalité de notre démocratie. Notre ambition est qu'elles soient avant tout des services publics vecteurs de citoyenneté et de lien social et qu'elles garantissent, grâce au principe de libre constitution des fonds, l'accès pluraliste à l'information et aux savoirs. Les bibliothèques se trouvent aujourd'hui confrontées aux enjeux du numérique dans les domaines culturels, pédagogiques et de recherche. Elles doivent également élargir leurs horaires d'ouverture et renforcer la médiation, afin de mieux répondre aux attentes et pratiques nouvelles de leurs publics et de contribuer encore plus à la réussite scolaire et universitaire. Elles sont également amenées à s'adapter aux configurations nouvelles de leurs territoires d'intervention, dans les sites universitaires comme dans les collectivités territoriales, dans un contexte de maîtrise des dépenses publiques.

L'inspection générale des bibliothèques (en liaison, le cas échéant, avec les autres inspections générales de nos deux départements ministériels) contribue à mesurer l'efficacité des politiques publiques en ce domaine et à accompagner les projets et les décisions de leurs différents acteurs. Ainsi, dans le cadre de cette lettre de mission, les rapports consécutifs aux missions et au contrôle technique des sites et des établissements et aux études thématiques devront proposer des recommandations concrètes et des pistes d'innovation concernant l'optimisation des moyens mobilisés au service des usagers. Vous vous attacherez en particulier aux améliorations à apporter pour adapter les horaires d'ouverture au rythme de vie des usagers, pour rendre encore plus efficientes l'organisation du travail et les politiques d'acquisition et de conservation de documents physiques et numériques, ainsi que pour renforcer la coopération entre les différentes catégories de bibliothèques dans les secteurs de l'enseignement, de la culture et de la recherche.

Les recteurs-chanceliers, les directeurs d'administration centrale ou les responsables des établissements publics qui souhaitent une intervention de l'inspection générale des bibliothèques sont invités à prendre l'attache du Cabinet du ministre concerné. Dans les mêmes conditions, l'inspection générale des bibliothèques est également susceptible d'intervenir pour le compte des collectivités territoriales qui en font la demande.

Nous vous prions d'agréer, monsieur le doyen, l'assurance de notre meilleure considération.

Najat Vallaud-Belkacem

Audrey Azoulay

Monsieur Pierre CARBONE
Doyen de l'Inspection générale des Bibliothèques
110 rue de Grenelle
75357 PARIS 07 SP

18 OCT. 2016

**Liste des missions**

**Missions pour le ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche :**

*** Sites et établissements :**

- le réseau documentaire des écoles françaises à l'étranger,
- la COMUE Léonard de Vinci et la documentation,
- Sorbonne universités et la documentation,
- le SCD de l'université des Antilles,
- le SCD de l'université de Guyane,
- le SCD de l'université de Caen,
- le SCD de l'université de Cergy-Pontoise,
- le SCD de l'université de La Réunion.

*** Missions thématiques :**

- les enjeux du numérique au regard des ressources documentaires en sciences humaines et sociales (numérisation, indexation, nouveaux usages) (cf. mesure n° 6 du plan en faveur des sciences humaines et sociales),
- le réseau des centres de formation aux carrières des bibliothèques,
- les nouveaux modes de coopération entre l'agence bibliographique de l'enseignement supérieur et les établissements dans le cadre du système de gestion de bibliothèque mutualisé (SGBM).

**Mission conjointe avec l'IGAENR**

Parcours et carrière des conservateurs de bibliothèque.

**Mission commune au ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche et au ministère de la Culture et de la Communication**

Suivi du plan d'accompagnement des éditeurs scientifiques pour le passage au numérique (mission conjointe IGAENR/IGAC/IGB).

**Missions pour le ministère de la Culture et de la Communication**

*** Bibliothèques municipales ou intercommunales :**

- Avignon,
- Bondy,
- Mulhouse,
- Rodez,
- Soissons,
- Vannes.

*** Bibliothèques départementales :**

- Allier,
- Doubs,
- Lot-et-Garonne,
- Maine-et-Loire,
- Nièvre,
- Yvelines.

*** Ville de Paris :**

- bibliothèque de l'Hôtel de Ville.

*** Établissements publics :** la politique documentaire de la Cité de l'architecture et du patrimoine.

# Annexe 2

**Personnes rencontrées**

Madame Elydia BARRET, responsable projet Archives, Grand Equipement Documentaire, Campus Condorcet
Madame Odile CONTAT, InSHS-CNRS
Monsieur Marin DACOS, directeur d'OpenEdition, CNRS
Madame Julie DEMANGE, archiviste, Bibliothèque de documentation internationale contemporaine
Madame Nathalie FARGIER, directrice UMS Persée
Madame Stéphanie GROUDIEV, chef de projet Grand Equipement Documentaire, Campus Condorcet
Monsieur Johann HOLLAND, chargé de mission Campus Numérique, Campus Condorcet
Madame Marie-Véronique LEROI, Département de l'innovation numérique, Ministère de la culture et de la communication
Madame Anaïg MAHE, maître de conférence, Urfist de Paris
Monsieur Marc MINON, directeur général de Cairn.info
Monsieur Thomas PARISOT, responsable relations institutionnelles de Cairn.info
Monsieur Stéphane POUYLLAU, directeur technique d'Huma-Num
Madame Valérie TESNIERE, directrice Bibliothèque de documentation internationale contemporaine