

Têtes chercheuses et archives à l'ère de l'informatique

Margot Georges, Magali Moysan

Citer ce document / Cite this document :

Georges Margot, Moysan Magali. Têtes chercheuses et archives à l'ère de l'informatique. In: La Gazette des archives, n°245, 2017. Meta/morphoses. Les archives bouillonnent de culture numérique – Forum des archivistes, 30-31 mars et 1er avril 2016. pp. 191-201;

doi : <https://doi.org/10.3406/gazar.2017.5526>

https://www.persee.fr/doc/gazar_0016-5522_2017_num_245_1_5526

Fichier pdf généré le 07/01/2020

Têtes chercheuses et archives à l'ère de l'informatique

Margot GEORGES

Magalie MOYSAN

Introduction

À la fin des années 1980, Tim Berners-Lee pose les principes du World Wide Web au sein du CERN (organisation européenne de recherche sur le nucléaire) dans l'ambition de connecter les physiciens du monde entre eux. Comme en témoigne cet exemple, l'informatique et la recherche ont une histoire commune. Nous nous sommes donc intéressées aux liens entre la recherche scientifique, l'informatique et les archives. Malgré une évidente polysémie, nous avons décidé d'étendre la notion d'informatique au numérique et à l'électronique, en la considérant dans son sens usuel servant à désigner tout ce qui touche l'utilisation des nouvelles technologies liées à l'ordinateur ou à Internet.

Cet article s'inscrit dans le contexte de deux doctorats en archivistique préparés à l'université d'Angers : l'un porte sur la place des archives dans le processus de recherche en sciences médicales de 1968 à 2006¹. L'autre étudie les représentations sociales des archives pour les chercheurs en sciences du végétal². Ces deux projets de recherche interrogent les relations que les chercheurs entretiennent avec les archives – notamment les leurs –, la définition qu'ils en donnent et la place qu'ils leur accordent dans leur travail.

¹ MOYSAN (Magalie), *La place des archives dans la recherche médicale en France entre 1968 et 2006*, sous la direction de Patrice Marcilloux.

² GEORGES (Margot), *Les chercheurs et leurs archives : étude des représentations dans les laboratoires des sciences du végétal*, sous la direction de Patrice Marcilloux et Bénédicte Grailles.

Ces travaux présentent des distinctions dans leur périmètre, leur approche et leur méthodologie. D'une part, ils étudient deux disciplines différentes, avec des matériaux et des pratiques de travail qui leur sont propres : les sciences médicales sont plus encadrées par la loi – y compris dans la conservation de documents – et leurs activités reposent depuis longtemps sur la mise en place et l'utilisation de bases de données. Les sciences du végétal, quant à elles, font parfois appel à l'histoire, et les contraintes légales ou l'utilisation de bases de données sont très variables. D'autre part, l'un de ces travaux s'adresse au responsable d'une structure de recherche qui est amené à se positionner pour le collectif, l'autre au chercheur en tant qu'individu. Enfin, l'étude sur les sciences médicales s'inscrit dans le passé et cible des structures créées avant 2001 pour analyser leurs pratiques entre 1968 et 2006, celle sur les sciences du végétal s'intéresse à la période actuelle pour étudier les représentations.

Dans cet article, nous essayerons de comprendre la façon dont l'informatique a fait évoluer les pratiques de production et de conservation des documents et données, ses impacts sur les échanges entre chercheurs et sur leur vision des archives. Notre réflexion s'appuie sur les résultats de deux enquêtes sous la forme de questionnaires adressés par Internet respectivement à des chercheurs en sciences du végétal (selon une définition large qui exclut cependant les ingénieurs et personnels techniques ou administratifs) et à des directeurs de structures de recherche en sciences médicales : les registres de morbidité, qui recensent tous les cas d'une pathologie sur un territoire donné, et les études de cohorte, qui suivent dans le temps un ensemble de sujets ayant vécu une même expérience¹. Le premier questionnaire est diffusé depuis décembre 2015 auprès d'un millier de chercheurs en Anjou et en Île-de-France et a recueilli 99 réponses exploitables, soit 10 % de la population cible. La seconde enquête est administrée depuis février 2016. Elle a recueilli 16 réponses de registres de morbidité (29 % de la population cible) et 12 réponses d'études de cohorte (22 % de la population cible).

L'informatique dans la recherche scientifique

En 1991, Jean-Marc Chaduc explique que dans l'administration française, « l'introduction de l'informatique dans la vie professionnelle date de trente ans,

¹ LECLERC (Annette), PAPOZ (Laure), BREART (Gérard), LELLOUCH (Joseph), *Dictionnaire d'épidémiologie*, Paris, Frison-Roche, 1990, p. 26.

celle de la micro-informatique de dix ans. Affaire de spécialistes au départ [...] elle est devenue plus récemment culture »¹. Dans le domaine scientifique, ce mouvement a commencé plus tôt puisque l'informatique s'est développée en lien avec la recherche et de nombreux outils ont été pensés ou conçus pour et par les scientifiques. Manuel Castells, qui souligne le développement conjoint de la recherche en génétique et de l'informatique, note par exemple que « la convergence technologique accentue toujours davantage l'interdépendance entre les révolutions biologique et micro-électronique, sur un plan matériel tant que méthodologique »². Pourtant, des inégalités perdurent face à l'outil informatique.

La place du numérique dans le quotidien du chercheur est encore peu étudiée par la sociologie des sciences, qui l'a surtout abordée à travers le prisme de l'inscription³. Deux dimensions nous ont donc paru intéressantes à observer : d'un point de vue individuel, les relations que les chercheurs entretiennent avec l'outil informatique, et d'un point de vue collectif, la place de l'informatique dans l'activité de recherche.

En sciences du végétal, les chercheurs maîtrisent l'outil informatique : 90 % d'entre eux utilisent un ordinateur à titre personnel tous les jours ou presque, 91 % se connectent à Internet tous les jours ou presque, 100 % ont déjà envoyé ou reçu des courriels et 84 % font des recherches documentaires en ligne. Dans leur travail, ils ont la sensation d'avoir des documents mieux rangés sur informatique qu'en papier et surtout d'y retrouver plus facilement une information. En revanche, malgré leur « double origine universitaire »⁴, les réseaux sociaux sont moyennement connus et peu utilisés par les chercheurs. En sciences médicales, l'avancement actuel de l'enquête ne permet pas de dresser un panorama précis ; notons cependant que les registres et les cohortes sont entièrement, et depuis longtemps, constitués autour de bases de données informatisées.

¹ CHADUC (Jean-Marie), « L'informatique, levier du changement dans l'organisation », *Réseaux*, n° 49, 1991, p. 121.

² CASTELLS (Manuel), *La société en réseaux. L'ère de l'information*, Paris, Fayard, 1998, p. 89.

³ LATOUR (Bruno), WOOLGAR (Steve), *La vie de laboratoire : la production des faits scientifiques*, Paris, La Découverte, 2006, 308 p.

⁴ DAGIRAL (Éric), « L'articulation des sociabilités et des savoirs en tension : les réseaux sociaux et l'université », *Transversalités*, n° 116, 2010, p. 31. Dans cet article, l'auteur explique que « d'une part, les modes de communication et la diversité des outils collaboratifs en ligne prennent leur essor dans les pratiques de chercheurs des campus américains au cours des années 1970, avant de s'étendre à d'autres catégories de population à partir des années 1990. D'autre part, bon nombre de sites dits de "réseaux sociaux" s'avèrent être le fait d'étudiants désireux d'intensifier les liens entre pairs ».

L'outil informatique s'est aussi largement diffusé dans la pratique scientifique. En sciences du végétal comme en sciences médicales, la gestion reste mixte entre papier et électronique, mais l'informatisation des activités est vécue comme complète. Dans le domaine du végétal, tous les chercheurs considèrent que leur activité est informatisée et qu'ils produisent des données. L'outil informatique est très largement représenté dans les lieux de conservation des données, documents et archives : l'ordinateur professionnel arrive en tête et le bureau comme lieu de travail en deuxième position.

Dans les sciences médicales, tous les registres et études de cohortes gèrent des documents et données électroniques avant 2006. Cependant, les volumes conservés varient beaucoup d'une structure à l'autre, allant du gigaoctet au téraoctet. Ces différences peuvent être mises sur le compte de stratégies de conservation différentes, certaines entités ayant mis en place une politique de numérisation systématique de leurs documents. Pour autant, 90 % des structures conservent plus d'un mètre cube d'archives papier.

Les pratiques documentaires des chercheurs dans l'environnement informatique

La recherche en sciences du vivant est donc très fortement informatisée et les chercheurs ont un rapport plutôt confiant à l'informatique. Celle-ci, en tant que dispositif technique, a même fait évoluer les usages et les rapports sociaux¹ dans la recherche : en sciences du végétal, 64 % des chercheurs indiquent avoir changé leur manière d'organiser leurs documents et leur archivage durant leur carrière. La généralisation de l'informatique est la première raison invoquée pour justifier ce changement. D'un point de vue documentaire, cette évolution a fait émerger plus fortement des notions comme le partage d'informations ou la sécurité des données personnelles.

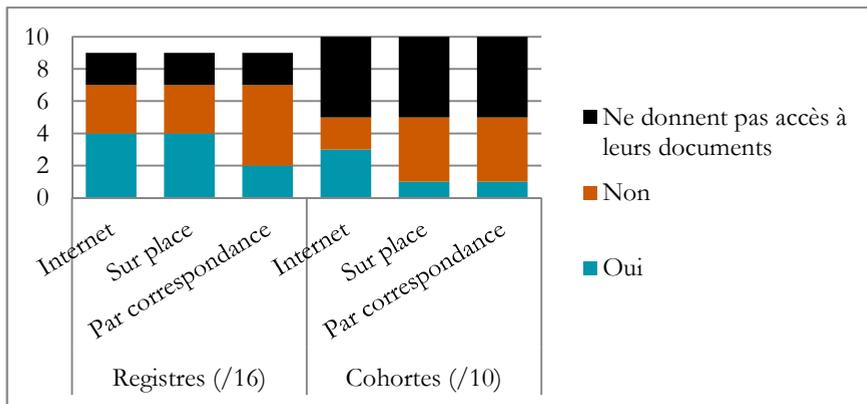
Échanger et partager

L'échange d'informations constitue l'un des piliers de la recherche scientifique et l'informatique puis Internet, en tant que médias, ont fait évoluer les modes d'échanges. En sciences médicales, 75 % des responsables d'études de cohortes

¹ ZARCA (Bernard), *L'univers des mathématiciens. L'éthos professionnel des plus rigoureux des scientifiques*, Rennes, Presses universitaires de Rennes, 2012, p. 41.

et 68 % des responsables de registres estiment que l'informatique a favorisé les échanges de documents et données. Ils précisent leurs réponses en mettant en avant des notions associées au gain de temps (« rapidité d'échanges », « gain d'efficacité ») et à la facilité (« facilité d'accès », « facilité d'archivage »). Ces notions peuvent aussi bien être associées à la collecte des données (« accessibilité aux données administratives », « facilité de croisement ») qu'à la gestion de leurs documents (« facilité d'archivage »). Les termes choisis indiquent une continuité dans les usages. Ainsi, pour les études de cohorte et les registres de morbidité, l'informatique ne semble pas avoir modifié les réseaux d'échanges, les sources et les interlocuteurs restant les mêmes. La différence réside davantage dans la réduction du temps accordé à certaines tâches et dans le gain de performance. Pour autant, parmi les répondants, deux registres relativisent l'apport de l'informatique dans le partage de données et soulignent une grande inégalité dans l'informatisation de leurs sources : un responsable de registre estime ainsi que seules 20 % des sources sont informatisées, et que « les données sources clés sont rarement informatisées ». Les usages varient donc selon le contexte et l'informatique n'a pas apporté de « règles sociales [de communication] unifiantes et unifiées » en 2006 comme elle a pu le faire dans le domaine de la physique¹.

Les dispositifs d'accès se tournent aussi largement vers l'utilisation de l'informatique, comme en témoigne le graphique ci-dessous. Avant 2006, Internet constituait déjà un moyen privilégié pour le partage de documents et données², avec, pour les registres, la consultation sur place.



Modes d'accès aux documents et données © Magalie Moysan

¹ DE LA VEGA (Josette), *La communication scientifique à l'épreuve d'Internet. L'émergence d'un nouveau modèle*, Villeurbanne, Presses de l'ENSSIB, 2000, p. 219.

² Hors bases de données et publications.

Sauvegarder, conserver et protéger

La protection des données s'inscrit comme le corollaire du partage chez les chercheurs en sciences médicales. Une partie des registres de morbidité et les études de cohortes conservent des données à caractère personnel et la sécurité de ces données est une préoccupation forte. Ainsi, de nombreuses mesures sont prises tout au long du processus documentaire : lors de la collecte des données (« déconnexion automatique en cas d'erreurs de saisie »), dans l'utilisation des données (charte utilisateur, restrictions d'utilisation), mais surtout dans la conservation : mots de passe, « stockage des disques durs dans un coffre ignifuge fermé à clef », « cryptage lors des envois », distinction de supports de conservation pour les données nominatives, etc. Le besoin de sécurité transparait également dans la conservation des archives papier : cinq registres indiquent ainsi disposer de salles d'archives sécurisées. Nous pouvons poser l'hypothèse que la loi CNIL a augmenté l'intérêt porté à la sécurité, y compris pour le papier, mais l'avancement actuel de nos travaux ne nous permet pas d'étayer cette idée. En sciences du végétal, la notion de sécurité est moins intériorisée par les chercheurs car ils manipulent des données moins confidentielles, donc soumises à un encadrement légal moins strict. Pourtant, l'action d'archiver, que ce soit en papier ou en électronique, est considérée comme proche de celle de « protéger ».

Cette action d'archiver fait l'objet d'un soin particulier dans l'environnement informatique : en sciences médicales, la quasi-totalité des structures observées ont mis en place des moyens spécifiques pour la conservation de leurs bases de données et documents électroniques sur le long terme, démontrant un réel intérêt pour ces enjeux. Celui-ci se traduit dans le choix des supports de conservation (disquettes, cédéroms, disque dur externes, zip, bandes LTO), dans la temporalité des sauvegardes (sauvegardes automatiques, sauvegardes quotidiennes ou « régulières »), dans la duplication (plusieurs copies, plusieurs lieux de conservation) et dans les acteurs mobilisés (deux registres font mention d'équipes informatiques dédiées). La sauvegarde des documents passe également par le changement de support : 19 % des registres et 41 % des études de cohorte avaient mis en place un système de numérisation avant 2006. Mais la numérisation de masse reste minoritaire. En sciences du végétal, où seules les pratiques individuelles des chercheurs ont été interrogées, la dématérialisation existe tout autant que la rematérialisation : 95 % des chercheurs impriment et 83 % d'entre eux numérisent. Ce sont des pratiques individuelles très répandues. Cependant, la nature des documents imprimés ou numérisés est différente. Dans les premiers nous retrouvons la documentation,

les projets d'articles, les contrats et conventions. Ce sont des documents utilisés pour le travail quotidien ou, dans une moindre mesure, pour leur valeur de preuve. Dans les seconds, les chercheurs citent les documents personnels, les rapports et à nouveau les documents probants, témoignant d'une numérisation axée sur la conservation.

Le classement des documents fait aussi l'objet d'une grande attention chez les chercheurs. Ainsi, tous les registres déclarent avoir mis en place un système de classement papier, quand ils sont 62 % à en avoir un en électronique. Dans les cohortes, le besoin de classement est moins ressenti mais on retrouve les mêmes proportions : les deux tiers ont un système papier et la moitié a un système pour l'électronique.

Enfin, ces pratiques de conservation sont focalisées sur l'usage. Une étude auprès des chercheurs en sciences humaines et sociales de l'université de Lille 3 montrait une pratique individuelle de la conservation¹. Dans le cadre des registres et cohortes, elle semble être davantage une pratique collective. Cela s'explique probablement par le caractère central de la base de données dans ces structures. Il semble qu'il existe également une tension entre partage et protection, qui a évolué dans le temps. Le développement d'Internet a favorisé les échanges mais a aussi demandé la mise en œuvre de moyens de partage sécurisés comme le cryptage. Pour autant, un équilibre semble installé et l'analyse des réponses des chercheurs en sciences médicales montre une certaine homogénéité des discours. En sciences du végétal, l'impact de l'informatique sur ces notions ne peut être analysé. Notons tout de même que dans la définition du mot « archiver », et de manière encore plus prégnante dans celle des activités de l'archiviste, protéger ou sécuriser et rendre disponible ont des places quasiment équivalentes.

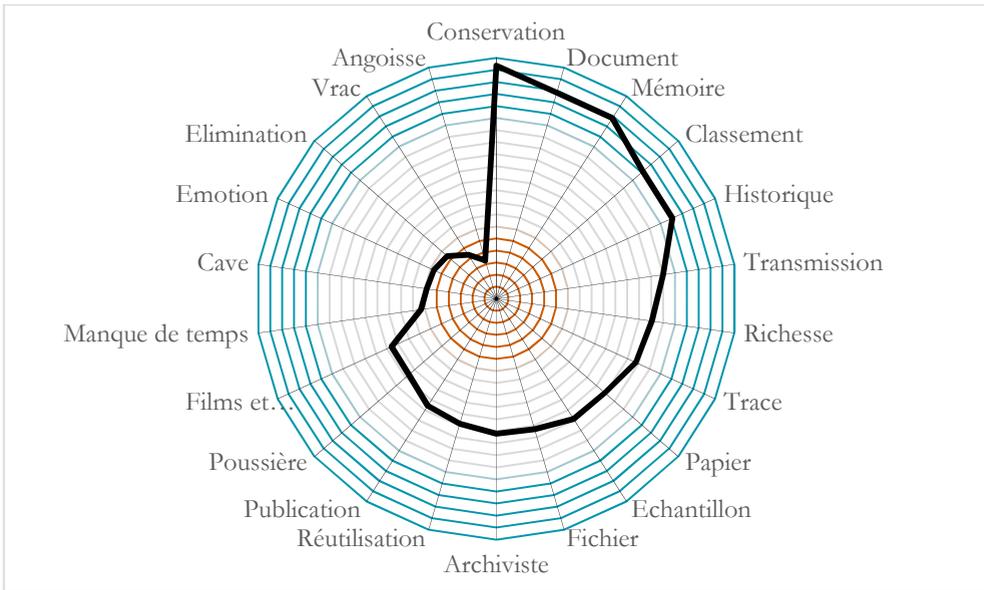
Les archives pour les chercheurs : quelle place pour l'informatique ?

Les chercheurs sont conscients de l'importance de conserver les documents dans l'environnement informatique. Pour autant, les considèrent-ils comme des archives ? Leur vision des archives prend-elle en compte l'informatique ?

¹ PROST (Hélène), SCHOPFEL (Joachim), *Les données de la recherche en SHS. Une enquête à l'Université de Lille 3. : Rapport final*, Rapport de recherche, Université Lille 3, 2015.

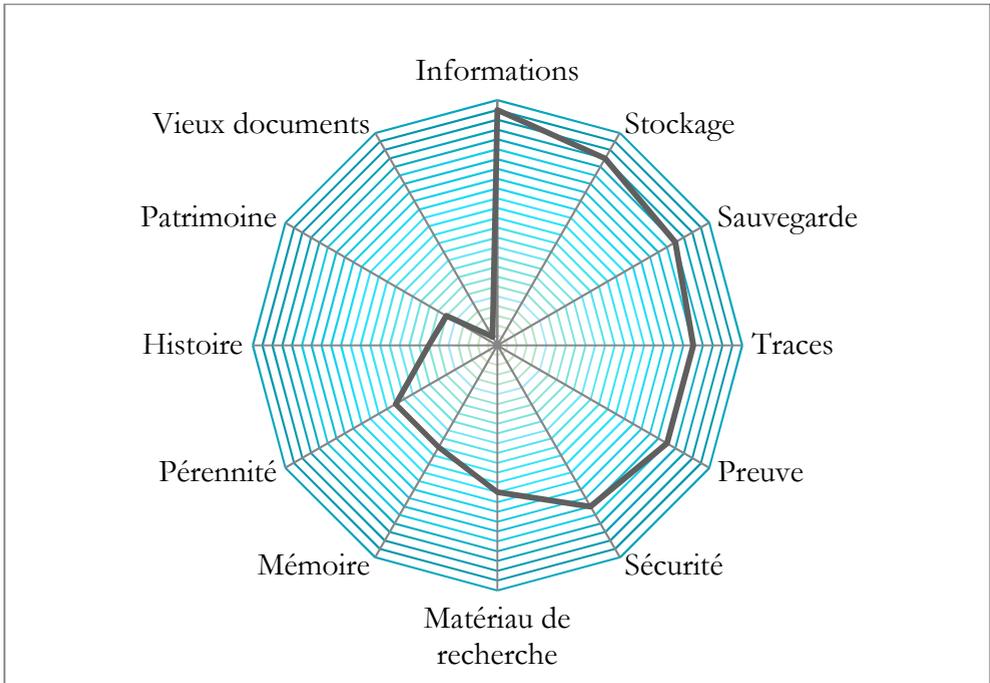
Archives et informatique

En France, l'année 2016 a été marquée par l'ajout d'une référence directe à l'informatique dans la définition des archives du Code du patrimoine¹. Pour les chercheurs, cette association ne va pas de soi. Ainsi, en sciences du végétal, sur 391 mots clefs recueillis en début de questionnaire, seuls 24 font référence à l'informatique (dont 11 aux données) contre 36 à l'information matérielle (dont 15 au papier). L'analyse des réponses orientées (schémas) témoigne également de l'absence d'association entre ces deux notions. Les chercheurs s'appuient plutôt sur leur expérience pour définir les archives : en sciences du végétal, les termes les plus cités sont « conservation », « document », « mémoire » et en sciences médicales « information », « stockage », « sauvegarde », l'usage influençant leur vision des archives.



Champ sémantique du mot « archives » en sciences du végétal © Margot Georges

¹ <http://archivistes.org/Loi-LCAP-I-AAF-a-fait-entendre-la-voix-des-archivistes>



Champ sémantique du mot « archives » en sciences médicales © Magalie Moysan

La notion de sauvegarde, très liée à l'informatique pour les chercheurs, se retrouve également dans la sémantique des expressions « archiver » et « activités de l'archiviste » en sciences du végétal. Nous nous sommes interrogées sur le sens de ce lien privilégié : les chercheurs ont-ils une plus grande conscience de l'importance d'archiver dans le contexte électronique ou pensent-ils la numérisation comme une forme d'archivage ? L'absence de proximité d'association des termes « numériser » et « archiver » par les chercheurs en sciences du végétal nous montre qu'ils font bien la distinction entre les deux. En outre, l'étude des lieux de conservation des documents, données et archives ne fait pas ressortir de migration des documents papier vers des supports informatiques pour l'archivage. D'autre part, les documents numériques sont bien plus souvent cités comme des documents devant être conservés par les chercheurs en sciences du végétal. En sciences médicales, l'association entre « archiver » et « numériser » est, elle aussi, plutôt à la marge ; seuls une cohorte et un registre font référence à la numérisation comme pratique d'archivage : « une numérisation rétroactive de toutes les archives papier du registre a été faite en 2011. Depuis, tous les documents papier sont

scannés et enregistrés, et détruits au fur et à mesure (chaque année à la fin de la validation d'une année d'enregistrement) »¹.

Le support des documents, papier ou informatique, n'est donc pas un élément essentiel dans la définition que les chercheurs en sciences du vivant donnent des archives. Leur vision de celles-ci dépend davantage de l'usage qu'ils en font et les chercheurs témoignent d'un souci de conserver les documents électroniques.

Archives et bases de données

En sciences du végétal, un répondant définit les archives comme « une base de données historique ». Pour finir, il nous a donc paru intéressant d'interroger ce lien entre archives et bases de données. En sciences du végétal, l'avancée actuelle de notre enquête ne permet pas d'en fournir une analyse. Par contre, en sciences médicales, les registres de morbidité sont constitués autour d'une base de données, propre à chacun et nourrie grâce à des documents sources, principalement collectés auprès de structures médicales. Cette base de données constitue la « mémoire » du registre. Pour autant, est-elle associée au terme « archives » ? La structuration autour d'une base de données rappelle celle étudiée par Lorna Heaton et Florence Millerand en botanique. Les deux chercheuses soulignent qu'une base de données « vise à rendre les données mobiles, c'est-à-dire indépendantes des environnements dans lesquels elles ont été produites, et pourtant pleines des précieuses informations documentaires qui permettront de les resituer dans leurs contextes d'origine »². La base de données s'inscrirait donc en opposition avec les archives, ancrées dans leur contexte de production, et se substituerait à celles-ci : « les données (et métadonnées) produites par le codage, l'enregistrement et la documentation, sur support informatique, des observations écologiques sont utilisées en lieu et place des matériaux d'origine, dont une partie seulement est conservée »³. Dans les registres de morbidité, la distinction entre la base de données et les archives apparaît également très nettement : aucun registre n'évoque sa base de données lorsqu'il décrit ses archives. Pourtant, celles-ci sont considérées comme des documents de valeur, puisque 87 % des registres jugent qu'il est essentiel de conserver les dossiers qui ont servi à alimenter la base de données. Elles remplissent alors une fonction de trace, pour permettre la recontextualisation.

¹ Réponse au questionnaire adressé aux responsables de registres de morbidité.

² HEATON (Lorna), MILLERAND (Florence), « La mise en bases de données de matériaux de recherche en botanique et en écologie. Spécimens, données et métadonnées », *Revue d'anthropologie des connaissances*, n° 7, 2013, p. 15.

³ *Ibid*, p. 16.

Conclusion

Les chercheurs ont pris part au développement de l'informatique et il existe des liens étroits entre ces deux entités. D'après nos données sur les chercheurs en sciences du vivant, nous pouvons confirmer que l'informatique est bien intégrée dans les pratiques de recherche. Elle a intensifié, modifié ou renforcé des pratiques d'échanges, de partage, de sauvegarde, de conservation et de protection de documents et données. Enfin, si elle a peu fait évoluer la définition que les chercheurs donnent des archives, elle a eu un impact réel sur celle qu'ils donnent de l'archivage.

Les chercheurs ont ainsi une vision des archives très centrée sur leurs usages, largement informatisés, ce qui explique qu'ils associent les termes « archivage » et informatique. Il existe des différences disciplinaires, qui influent sur les pratiques de gestion documentaire, comme l'existence de données personnelles ou la structuration de la recherche autour de bases de données. La poursuite de nos travaux de thèses nous permettra peut-être d'identifier d'autres facteurs d'impact sur la relation entre chercheurs, informatique et archives, tels que l'organisation du travail au sein de la structure de recherche, l'âge ou l'acculturation personnelle.

Margot GEORGES

Doctorante en archivistique à l'université d'Angers
et archiviste à l'École pratique des hautes études (EPHE)

Sujet de thèse : « Les chercheurs et leurs archives : étude des représentations dans les laboratoires des sciences du végétal »,
sous la direction de Patrice Marcilloux et Bénédicte Grailles

margot.georges@laposte.net

Magalie MOYSAN

Doctorante en archivistique à l'université d'Angers
et archiviste à l'université Paris Diderot

Sujet de thèse : « La place des archives dans la recherche médicale en France
entre 1968 et 2006 », sous la direction de Patrice Marcilloux

magalie.moysan@gmail.com