

# Enjeux de la numérisation des herbiers pour l'information et la communication scientifiques : de la transformation des matières documentaires à l'évolution des pratiques

*Digitalization of naturalist collections: improvements for scientific communication*

*Aportaciones de la numeración de colecciones de historia natural para su explotación científica*

Article inédit, mis en ligne le 27 octobre 2015

## Lisa Chupin

*Lisa Chupin est doctorante en sciences de l'information et de la communication au laboratoire Dicen-IDF (EA 7339) et ATER au CNAM Ile-de-France. Ses travaux croisent les problématiques de valorisation du patrimoine numérique, de participation des publics dans des démarches dites de crowdsourcing ou de sciences citoyennes, et d'exploitation scientifique des corpus numérisés (plus d'informations à l'adresse : <http://dicen-idf.org/membre/chupin-lisa/>).*

## Plan de l'article

Introduction

Fondements et renouvellements de la valeur documentaire des herbiers

*Enjeux de la numérisation des collections pour l'exploitation et la communication de l'information scientifique*

*L'herbier comme artefact médiateur de la production des savoirs*

*Le rôle de la constitution des collections dans la communication de l'information scientifique*

*Des spécimens dont la matière fait preuve*

Méthodologie de l'étude des pratiques mobilisant des herbiers physiques et numériques

*Modalités de l'observation et de l'analyse des pratiques mobilisant des herbiers*

*Le statut documentaire de l'image appréhendé par les pratiques informationnelles*

Les bases de données d'images d'herbiers, entre « amplification » et « réduction » de l'information scientifique

*Les limites de l'efficacité des bases de données pour la recherche de spécimens*

*L'accès à l'information scientifique facilité par la documentation des images*

La numérisation à l'origine de nouvelles modalités d'exploitation des collections

*Les limites d'une exploitation massive des données d'herbier*

*Le statut documentaire des images fonction des objectifs présidant à leur exploitation*

Communication médiatisée par les images et documentation des collections

*Les interactions entre usagers et gestionnaires médiatisées par les collections numérisées*

*Les images médiatisant le partage d'informations entre institutions*

Conclusion

Références bibliographiques

## Résumé

Nous appréhendons les mutations entraînées par la numérisation des collections d'histoire naturelle dans la production et la communication de l'information scientifique. Notre étude s'appuie sur une analyse des herbiers comme artefacts répondant aux besoins de la conservation et de la communication de l'information entre acteurs intervenant dans la production des savoirs mobilisant les collections. Nous analysons le statut documentaire des images à partir de leur exploitation différentielle dans les pratiques scientifiques selon les objectifs et étapes des recherches. Malgré les

limites de l'exploitation des bases de données d'images pour le travail sur les spécimens, elles sont essentielles à l'accès et à l'enrichissement des informations des collections.

### **Mots-clés**

Patrimoine numérisé, communication scientifique, artefact, herbier.

### **Abstract**

We analyze how the digitalization of naturalist collections improves their scientific exploitation. To this aim, we observed the working practices of naturalists that lead them to consult herbaria. We compare the information value of a digitalized specimen with that of the physical one with an analysis of the role of each medium. Referencing a physical document is mandatory when describing a new species, but novel uses of digital herbaria increase: they make searching in international collections easier and can be used for the identification of a specimen. The digital herbarium is the medium of a new documentation about collections, which is built by joint efforts of users.

### **Keywords**

Digital heritage, scientific communication, document, herbarium.

### **Resumen**

Analizamos las aportaciones de la numeración de colecciones de historia natural para su explotación científica a partir de la observación de las prácticas que motivan la consulta de herbarios. Comparamos la situación documental del espécimen analizado con respecto al de la colección física. Si el herbario físico sigue siendo una referencia obligada para describir una especie nueva, los usos de la base de datos se multiplican: además de la investigación facilitada en las colecciones, las imágenes permiten la identificación del espécimen y facilitan la producción colaborativa de documentación sobre las colecciones.

### **Palabras claves**

Numeración, patrimonio científico, documento, herbario.

### **Introduction**

La numérisation de fonds patrimoniaux conduit à envisager la « cyberscience » (Nentwich, 2003) sous l'angle des transformations des disciplines liées à l'exploitation de corpus numérisés. Ces mutations ont été abordées dans le domaine des humanités, notamment au niveau des pratiques de recherche d'information des historiens, et de leur rapport aux professionnels de la documentation (Vinck et Natale, 2015). Nous nous penchons sur les enjeux que représente pour la production de

savoirs scientifiques la numérisation des herbiers, collections de plantes séchées qui, en plus de leur valeur historique et patrimoniale, constituent une référence toujours valide et utilisée pour l'étude de la classification du vivant et de l'évolution des espèces. Les programmes de numérisation des herbiers ont été étudiés dans le champ des études des sciences et des technologies (STS) qui ont mis en lumière des modifications des pratiques de travail des chercheurs, dans le contexte de politiques scientifiques favorables à leur développement (Hine, 2008). La constitution de bases de données d'herbiers, marquée par les procédures de contrôle et de standardisation, a été étudiée comme une partie intégrante de l'activité scientifique, étant donné les recherches d'information qu'elle suppose (Heaton et Millerand, 2013). Nous questionnons les enjeux des bases de données issues de la numérisation des herbiers pour l'exploitation scientifique des collections. À partir d'une enquête sur les pratiques donnant lieu à la consultation des herbiers, réalisée auprès d'utilisateurs de collections appartenant à des institutions membres ou proches du réseau des collections naturalistes ReColNat (2013), nous cherchons à déterminer si l'image assume, dégrade, ou renouvelle les fonctions traditionnelles de l'herbier, non seulement de conservation, mais aussi d'échange d'informations entre utilisateurs des collections distants dans l'espace et le temps. Nous commençons par analyser les caractéristiques matérielles des collections d'herbiers qui leur permettent d'exercer leurs fonctions dans la production et la communication de l'information scientifique, avant d'envisager les conséquences de leur numérisation au niveau des pratiques informationnelles des acteurs de la production des savoirs – conservateurs, gestionnaires ou chercheurs, professionnels ou non.

## **Fondements et renouvellements de la valeur documentaire des herbiers**

### ***Des spécimens dont la matière fait preuve***

Comprendre le rôle de la matérialité de l'herbier dans ses fonctions de conservation et de communication de l'information scientifique est nécessaire pour prendre la mesure des enjeux de sa numérisation. Étant donné le rôle essentiel joué par l'étude des spécimens pour les disciplines naturalistes, l'herbier a servi de paradigme pour analyser le rôle de l'objet documentaire dans sa matérialité. Bruno Latour (1996) a décrit la production de l'information scientifique comme le transport vers un « centre » documentaire d'objets matériels capturés, collectés, ou dessinés sur le terrain, auquel ils restent toujours « branchés » par la « connectique » des inscriptions grâce auxquelles le spécimen étudié dans le centre est relié au terrain où il a été prélevé. La configuration de l'espace documentaire est aussi essentielle à l'exploitation du document, puisqu'elle permet la mise en série des spécimens pour les comparer entre eux : les informations dont chaque spécimen est porteur, malgré les « réductions » subies par rapport à l'observation sur le terrain, s'en trouvent « amplifiées ». La numérisation questionne la valeur documentaire de l'image qui ne conserve que la « forme » du spécimen et perd l'extrait de sa matière conservé dans l'herbier. Elle interroge la capacité de la base de données donnant accès aux images à assumer les fonctions d'amplification de l'information de chaque spécimen en offrant de nouvelles possibilités de « calcul » par la mise en relation de plusieurs images.

### ***Le rôle de la constitution des collections dans la communication de l'information scientifique***

Comprendre le rôle des collections dans la production des savoirs demande toutefois de compléter cette analyse de la chaîne de médiations matérielles (Jeanneret, 2009) par la prise en compte des opérations d'interprétation et d'appropriation de l'information qui en sont les corollaires. Si le traitement documentaire se traduit par une transformation matérielle du document, Marie Després-

Lonnet (2014, p.106) propose de le considérer comme un ensemble d'opérations de sélection et d'interprétation d'informations : ces dernières sont constitutives du processus de réduction et d'amplification de l'information des spécimens, mais aussi de la réalisation d'un projet de communication des savoirs, auquel répondent les collections physiques ou numériques (Després-Lonnet, 2000). Cette analyse des opérations intellectuelles nécessaires à la « documentation » des images des collections numérisées ouvre des perspectives pour comprendre les nouvelles amplifications de l'information qui peuvent en résulter. Considérer l'espace documentaire comme une réponse à des besoins de communication de l'information scientifique pose aussi la question de la capacité des bases de données d'images à assumer la même fonction, sachant que le classement dans l'espace physique des collections est porteur en lui-même d'indications sur l'interprétation à faire de la planche (Heaton et Millerand, 2013). Le rôle des collections dans la communication scientifique peut aussi être analysé au niveau de chaque spécimen à partir de ses caractéristiques documentaires.

### ***L'herbier comme artefact médiateur de la production des savoirs***

La notion d'« artefact », qui renvoie à l'intention de fabriquer un objet pour répondre à des besoins, s'applique bien à l'herbier dont le format s'est stabilisé parallèlement à la constitution des principales disciplines qui l'utilisent – systématique et écologie. Dans ces disciplines, l'accès à l'information scientifique passe par les publications autant que par les spécimens qui y sont mentionnés : la conservation du spécimen physique avec les expertises dont il a fait l'objet permet de vérifier ces dernières en réitérant l'analyse du spécimen. Symétriquement, toute publication décrivant une nouvelle espèce est assortie obligatoirement du dépôt dans une collection institutionnelle de spécimens ayant servi à l'étude, dits spécimens « types », avec l'obligation pour l'institution d'en assurer la communication sur demande. Resituer le document dans ce « processus de communication pour partie différé [...] entre des producteurs et des récepteurs liés par des intérêts communs » (Zacklad, 2004), intérêts qui sont en l'occurrence la contribution à la production des savoirs, permet de comprendre les spécificités documentaires des herbiers. Si l'étiquette atteste de la présence d'une espèce à un moment et à un lieu déterminé, cette première « documentarisation » (Zacklad, 2015), essentielle à l'exploitation de la planche d'herbier par des usagers séparés par de grandes distances spatiales et temporelles, se prolonge par l'ajout d'une étiquette chaque fois que la planche fait l'objet d'une nouvelle « détermination » – identification de l'espèce à laquelle appartient le spécimen (voir illustration 1). La transformation des documents dans leur matérialité par l'ajout d'étiquettes accompagne la production des savoirs à partir des spécimens. En ce sens, ce processus peut être lu comme un ensemble de « transactions » définies par Manuel Zacklad (2015) comme des « interactions productives », « permettant la transformation d'un artefact médiateur » – à savoir l'herbier – et « des personnes parties prenantes pour réaliser une performance », en l'occurrence les acteurs investis dans la production de savoirs botaniques. Cette conception interroge la capacité de l'image à servir d'« artefact médiateur » aux interactions entre acteurs de la production des savoirs, et celle de la base de données d'images à reproduire la « fluidité » (Zacklad, 2015) de la documentarisation des spécimens accompagnant les expertises successives dont ils sont l'objet.

### ***Enjeux de la numérisation des collections pour l'exploitation et la communication de l'information scientifique***

Envisager la matérialité du document non seulement comme support permettant la concentration et la confrontation de spécimens dans un même espace, mais aussi comme artefact qui répond aux besoins d'une communication scientifique largement distribuée permettant la capitalisation d'expertises dans les collections, conduit à envisager les enjeux de la numérisation des collections

pour la production des savoirs à différents niveaux. Plusieurs critères apparaissent pour établir s'il est légitime de considérer la base de données d'images comme un nouveau « centre » documentaire : sa capacité à accumuler et organiser des images pour faciliter leur confrontation, mais aussi celle de favoriser l'accès des usagers aux documents des collections ; en outre, reproduire la communication scientifique dont l'herbier est le support demande de rendre possible des « transactions » entre acteurs de la production des savoirs médiatisées par les images en leur donnant la possibilité de participer à leur documentarisation. Comparer le statut des bases de données d'images aux « centres » documentaires physiques requiert donc une analyse des modalités d'accès aux images et de leur exploitation, ainsi que des formes de contribution ouvertes aux usagers dans les bases de données.

## **Méthodologie de l'étude des pratiques mobilisant des herbiers physiques et numériques**

### ***Le statut documentaire de l'image appréhendé par les pratiques informationnelles***

Nous abordons la question du rôle des bases de données d'images d'herbiers dans l'exploitation et la communication de l'information scientifique à partir de l'étude des pratiques qui donnent lieu à leur consultation. Nous nous fondons sur une approche utilisée par Marie Després-Lonnet (2014) pour analyser le statut des photographies en archéologie, selon laquelle le document acquiert son statut dans les pratiques qui le mobilisent comme tel, faisant droit à l'intention de l'utilisateur qui peut se réapproprier l'image pour différentes finalités. La diversité des usages dont l'herbier fait l'objet selon les disciplines - 72 selon Vicky Funk (2003) - rend pertinente cette approche de la valeur informationnelle des herbiers selon les modalités plurielles de leur exploitation. Notre enquête a cherché à comprendre le rôle des spécimens, physiques et numériques, en resituant l'image parmi les différentes ressources mobilisées par les acteurs de la production des savoirs pour accéder à l'information scientifique et l'exploiter, selon leurs différents objectifs de recherche. Nos entretiens sont construits autour de la compréhension des pratiques informationnelles des chercheurs interrogés, entendues comme « manières dont l'ensemble des dispositifs, des sources, des outils, des compétences cognitives sont effectivement mobilisés dans les différentes situations de production, de recherche, traitement de l'information » (Chaudiron et Ihadjadene, 2010).

Dans la mesure où cette enquête questionne, entre autres documents mobilisés au sein d'une même activité, l'utilisation faite de l'image, elle recoupe en partie l'étude de Christine Hine (2013) concernant les conséquences de la numérisation des collections sur le travail des systématiciens. À partir de 22 réponses obtenues suite à un appel à témoignage envoyé sur une liste de diffusion internationale de systématique concernant le travail fait avec les images d'herbiers, elle a montré la prégnance d'une « culture matérielle », qui reste aujourd'hui marquée par la reproduction en ligne du travail sur l'herbier papier. Il n'est pas question de substituer les images aux collections existantes (Hine, 2008, p. 117) ; cependant, des pratiques faisant une place à l'image dans la recherche sont en émergence avec, outre l'accès facilité aux collections, l'identification de spécimens en ligne, mais aussi quelques expériences de sollicitation de l'expertise de chercheurs distants concernant la détermination de l'espèce d'un spécimen à partir de sa photographie. Si nos résultats concernant les principaux usages de l'image convergent avec cette étude, nous les analysons dans la perspective de la compréhension du rôle des bases de données comme « centres » documentaires capitalisant et mettant à disposition l'information scientifique, susceptibles de favoriser de nouvelles interactions productrices de savoirs.

### ***Modalités de l'observation et de l'analyse des pratiques mobilisant des herbiers***

Notre enquête a commencé en septembre 2013 dans le cadre de l'infrastructure de numérisation et de valorisation de collections naturalistes francophones e-ReColNat. Outre les synthèses des notes prises pendant les visites dans différents herbiers partenaires, et la participation à un colloque en janvier 2014 consacré au « rôle des collections botaniques dans la recherche sur les écosystèmes tropicaux » rassemblant des étudiants et des chercheurs au Cirad (Montpellier), la présente étude se base sur des entretiens d'une heure à une heure et demie, réalisés jusqu'en décembre 2014, avec 12 chercheurs engagés dans des problématiques liées à la systématique, à l'écologie ou à l'histoire des collections, ayant ou ayant eu des responsabilités dans la gestion d'herbiers. Les entretiens étaient axés sur l'explication du rôle des collections dans les activités des personnes interrogées, et ont donné lieu à des démonstrations commentées de la consultation des collections et des bases de données utilisées. Lors d'une journée de formation destinée aux gestionnaires d'herbiers dispensée par l'association *Tela Botanica*, en octobre 2014, nous avons également observé en situation des activités d'informatisation des collections avec recherche de spécimens « types ».

Cette enquête nous a conduit à analyser les dispositifs les plus fréquemment mentionnés par les usagers dans l'accès aux collections, physiques et numériques, à savoir « Sonnerat » (MNHN, 2013), la base de données des collections du Musée national d'Histoire naturelle (MNHN), et celle de « JSTOR Global Plants », portail international de spécimens et de publications numérisés. Nous avons également analysé comme une source d'information complémentaire pour comprendre l'activité en ligne suscitée par les collections numériques les commentaires laissés depuis trois ans (jusqu'en décembre 2014) relativement aux spécimens dans le site « *JSTOR Global Plants* ». Une première lecture d'une partie des commentaires nous en a montré le caractère fortement stéréotypé. À partir du corpus des commentaires extraits du site, nous avons complété ce premier constat par une analyse quantitative de la répartition des commentaires par spécimen et par commentateur.

### **Les bases de données d'images d'herbiers, entre « amplification » et « réduction » de l'information scientifique**

#### ***L'accès à l'information scientifique facilité par la documentation des images***

La fonction des bases de données la plus fréquemment mentionnée par les usagers d'herbiers que nous avons rencontrés est l'accès aux collections. Les bases de données d'images dotent les collections de catalogues : l'accès à l'information scientifique est facilité grâce à un nouveau mode de recherche de spécimens par requêtes dans les bases de données des institutions ayant mis leurs collections en ligne. Les usagers apprécient la plus grande autonomie qu'ils ont dans l'accès aux spécimens, et réciproquement, les gestionnaires disent recevoir des demandes plus précises de la part des usagers. La localisation dans différentes institutions des spécimens utiles à un chercheur - qui va par exemple tenter de réunir le maximum de spécimens d'une espèce qu'il étudie dans une région déterminée - est une partie intégrante de son travail, qui tire grand profit des bases de données. Ces dernières permettent de réunir plus rapidement un grand nombre de spécimens en vue de leur comparaison.

Les possibilités de comparaison de spécimens sont d'autant plus importantes que la numérisation est associée à la mise en réseau de collections. Il est alors possible de formuler des requêtes dans une unique base pour avoir accès à tous les spécimens indexés du réseau. C'est le cas avec la base de recherche du MNHN, « Sonnerat », qui a accueilli progressivement les données d'autres « Herbiers

de France », à l'origine du réseau des collections d'histoire naturelle ReColNat. On retrouve cette volonté de rassembler des données sur le site *JSTOR Global Plants*, qui est le premier outil cité dans notre enquête pour rechercher les spécimens « types », mais aussi les publications qu'il leur associe. Mais en dehors de ces initiatives de regroupements, de multiples institutions ont entrepris une numérisation en interne : la centralisation des données sur les collections reste donc partielle, malgré un catalogue recensant les collections (et non leur contenu précis) au niveau mondial, l'*Index Herbariorum* (2007). Il en résulte des pratiques de recherche de spécimens dans de multiples bases selon les besoins, la recherche dans de multiples « centres » accroissant les possibilités d'amplification internes à chaque collection.

Cet accès facilité aux collections tient non seulement aux possibilités de requêtes informatiques, mais aussi au travail de documentation des images, expression que nous employons pour désigner les opérations de déchiffrement, contextualisation, et recherche d'information nécessaires à produire une indexation fine à partir de la transcription des différentes informations présentes dans les étiquettes, et à lier le spécimen à des ressources associées (publications, carte géolocalisant le lieu de sa récolte). La valeur ajoutée de ce travail ressort de l'analyse des multiples recherches qu'il nécessite, mises en évidence par Serge Proulx et Lorna Heaton (2012) à partir de l'observation de la recherche des types dans l'herbier de l'université de Montpellier : les auteurs montrent que la numérisation à proprement parler (c'est-à-dire la photographie du spécimen) n'est que la partie la plus répétitive d'une activité plus complexe demandant la consultation de différentes ressources pour déterminer le statut de type d'un document. Cette conclusion ressort aussi de nos observations au Jardin Botanique de Lyon de la réalisation concomitante de la numérisation des types avec un travail de révision et d'inventaire de la collection, dont la numérisation de spécimens choisis est l'aboutissement. L'inventaire de la collection accompagné de sa révision donne lieu à une recherche sur la validité du nom associé à l'espèce de chaque spécimen, ce qui facilite pour l'utilisateur la recherche dans la base, qu'il peut interroger à partir des termes faisant actuellement consensus. Ce travail terminologique, dont la base de données capitalise les résultats, fait partie intégrante des activités des chercheurs, dont il peut être une étape nécessaire et même une spécialité de recherche, soit parce qu'il n'a pas été encore réalisé pour les spécimens auxquels s'intéresse le chercheur, soit à titre de vérification des informations fournies par la base. L'étude des recherches d'information préalables à la numérisation permet de comprendre que, outre l'accès à l'image, c'est tout autant le catalogue en ligne des collections qui facilite l'accès à l'information scientifique. Les lacunes de la documentation des images peuvent d'ailleurs en freiner l'exploitation en ligne.

### ***Les limites de l'efficacité des bases de données pour la recherche de spécimens***

Même dans le cas des processus de numérisation associés à des opérations de documentation des images, toutes les informations inscrites sur une planche sont rarement transcrites. Quand un procédé de « numérisation industrielle » (Pérez et Pignal, 2013) est choisi, comme celui appliqué pour l'herbier du MNHN qui a fait l'économie des ressources nécessaires à opérer la sélection des types pour rendre accessible un plus grand nombre d'images, il ne s'accompagne pas d'une documentation fine de chaque planche. La numérisation crée alors un nouveau travail de recherche de spécimens par tri des résultats de requêtes dans les bases de données d'images, qui remplace la recherche dans les liasses de planches, comme un nouveau préalable indispensable à leur exploitation.

Les bases de données ne sont d'ailleurs pas une entrée dans les collections privilégiée par tous les chercheurs interrogés. Les conservateurs affirment leur rôle dans la proposition aux visiteurs d'autres spécimens que ceux trouvés avec la base, à partir de leur connaissance des collections. Certains usagers insistent sur l'accès aux documents physiques comme une entrée d'autant plus importante

dans les collections qu'elle permet d'autres fonctions que le simple accès au document. Ainsi, parcourir les « compactus » (casiers empilés pour le classement des planches, équivalents des rayonnages de livres) pour un besoin précis, peut en fait amener à s'arrêter sur une anomalie de classement : simple erreur, elle peut aussi mettre sur la piste d'une découverte. Ils reconnaissent toutefois le caractère indispensable des bases de données des collections importantes au plan international auxquelles ils n'ont pas facilement accès, en particulier pour chercher la bibliographie associée à une espèce et les spécimens types correspondants.

L'usage largement partagé des bases de données d'herbiers pour accéder aux informations de multiples collections permet de les considérer comme de nouveaux « centres » documentaires dans l'environnement informationnel botanique. Si elles favorisent l'amplification de l'information des spécimens par leur comparaison, elles sont utilisées en complémentarité des collections physiques, dans une mesure qui varie en fonction des moments et des objectifs de la démarche scientifique, plus ou moins affectés par les réductions de la numérisation.

## **La numérisation à l'origine de nouvelles modalités d'exploitation des collections**

### ***Le statut documentaire des images fonction des objectifs présidant à leur exploitation***

Nous avons constaté chez la plupart des usagers rencontrés la limitation de l'utilisation de la base à un outil d'accès au document physique, qui reste le véritable support de l'analyse. Cela s'explique par l'impossibilité de procéder aux examens nécessaires à répondre à certaines questions de recherche à partir de l'image. Cette dernière ne suffit plus quand la matière doit elle-même être transformée pour procéder aux examens, par exemple quand il faut d'abord réhydrater un spécimen et en disséquer une partie, pour en mesurer un élément. Aussi, les activités prélevant directement des informations de l'image pour les traiter sont celles qui sont centrées sur la lecture des étiquettes, avec par exemple pour but de retrouver les spécimens de la collection d'un botaniste répartis entre différentes institutions. Mais il est aussi possible de travailler avec l'image quand les caractères observables au zoom suffisent.

L'exploitation de l'information de l'image pose la question de son statut par rapport à celui du document physique. Nous avons constaté la reconnaissance de la valeur de l'image comme aide à la détermination de l'espèce de nouveaux spécimens collectés, destinés à rejoindre les collections physiques. L'identification peut alors suivre la même démarche d'observation de différents caractères morphologiques significatifs que celle qui se fonde sur l'examen d'une planche d'herbier physique ; l'observation est même facilitée par les fonctions de zoom. Cette pratique d'identification au moyen d'images en ligne du spécimen physique destiné à rejoindre les collections nuance l'idée d'une dégradation des informations de l'image, qui entraînerait un statut nécessairement subordonné au document physique : le rapport est ici renversé puisque l'image devient un document de référence pour l'identification de nouveaux spécimens adjoints à la collection.

### ***Les limites d'une exploitation massive des données d'herbier***

Nous avons également été témoin (notamment lors de débats suscités lors de la journée d'étude de janvier 2014 au Cirad) d'une réflexion dans la communauté scientifique autour des précautions à prendre pour exploiter les données des bases d'images : la base de données permet de générer automatiquement à partir d'un jeu de données des cartes de répartition par exportation de données de géolocalisation de spécimens, documents traditionnellement obtenus par compilation de données

d'observations ou de collections. L'idée d'une synthèse automatisée des informations sur la répartition d'une espèce rencontre l'obstacle d'un nécessaire examen critique des documents, préalable à l'export des données, les erreurs de transcription lors de l'informatisation pouvant s'ajouter à d'autres déjà présentes dans les collections physiques. Si la possibilité de récupérer un jeu de données géolocalisées et d'en visualiser automatiquement la répartition est appréciée, elle demande en amont le même examen critique que celui appliqué lors d'un inventaire de données recherchées dans les collections physiques : la détermination de chaque spécimen est vérifiée, ainsi que la cohérence des données géographiques avec l'état actuel des connaissances de la répartition de l'espèce étudiée. L'exploitation des données issues de la documentation des images se fait donc dans la continuité des démarches d'examen des spécimens conservés sur papier. Elle ne permet de réaliser qu'une partie des opérations dont les spécimens sont le support, principalement la recherche dans les collections, qui se voit renouvelée par les bases de données. La communication de l'information scientifique reste imparfaitement reproduite quand elle passe par l'accès aux images, car les inscriptions issues de nouvelles analyses ne peuvent être apposées que sur un spécimen physique. Toutefois, l'enrichissement des données associées à l'image permet de reconstituer partiellement la structure transactionnelle de la communication scientifique en botanique.

## **Communication médiatisée par les images et documentation des collections**

La recherche d'informations sur les spécimens en vue de la documentation des collections en ligne suscite de nouvelles formes de communication médiatisée par les images, qui se traduisent par l'enrichissement des bases de données d'herbiers, reconstituant une possibilité de documentariser le document en ligne suite à son examen.

### ***Les images médiatisant le partage d'informations entre institutions***

Les collections numériques sont un outil très efficace pour faire progresser la documentation des spécimens numérisés, qui demande des recherches dans plusieurs collections. En effet, une image précisément indexée peut être utilisée pour documenter une planche qui l'est moins, à partir du moment où il est établi que les deux spécimens proviennent de la même récolte. Les recherches dans les bases de différentes institutions sont d'autant plus fécondes que les collections se constituent par agrégation de différents herbiers, détenus par des botanistes ayant eux-mêmes procédé de leur vivant à de nombreux échanges de spécimens avec leurs pairs (Pérez et Pignal, 2013).

Les facilités d'accès aux ressources des collections grâce aux images sont à l'origine d'une reconfiguration des pratiques de recherche des spécimens « types » qu'elles contiennent. Lors de la journée de formation à la recherche de spécimens types organisée par *Tela Botanica* dans le but de diffuser les pratiques développées dans le cadre du projet *JSTOR Global Plants* au profit de la connaissance du patrimoine naturaliste, il est apparu que cette activité était facilitée par le fruit des travaux antérieurs qui ont conduit à la mise en ligne d'un nombre important de spécimens types et de ressources utiles à leur recherche. D'après le témoignage d'une conservatrice d'un herbier régional, les pratiques de recherche de types dans les collections, antérieurement au projet *Global Plants Initiative*, prenaient pour appui les publications des botanistes représentés de manière importante dans la collection : ces ouvrages attestaient des spécimens types qui avaient été déposés lors de la publication d'une description d'espèce, susceptibles d'être toujours présents dans la collection. Aujourd'hui, on utilise en premier lieu les ressources issues des programmes de numérisation déjà menés à bien, c'est-à-dire les sites rassemblant déjà des spécimens types, à commencer par le site *JSTOR Global Plants*, et les bases bibliographiques pour trouver les publications ayant servi à décrire

chaque espèce, dits « protologues ». On peut considérer le processus de documentation des collections entrepris simultanément dans différentes institutions comme un ensemble de « transactions », médiatisées par les images, entre des gestionnaires qui consultent des ressources que d'autres ont déjà mis en ligne, afin de documenter les images des collections dont ils ont la responsabilité. Le mouvement de numérisation des collections entrepris par la plupart des grandes institutions naturalistes donne lieu à une progression concomitante de la documentation des collections, de nouveaux « centres » se constituant grâce à l'information rendue plus facilement accessible par les collections déjà numérisées. Les collections déjà en ligne, en facilitant l'accès à l'information, favorisent ces échanges entre institutions, qui restent toutefois en deçà du travail en réseau des bibliothèques : cela s'explique notamment par le fait qu'il n'y ait pas de notice commune à plusieurs planches qui puissent être échangées comme celles des livres. Les efforts de valorisation des collections d'histoire naturelle tendant à exploiter ces échanges au service de la connaissance des collections se traduisent par la recherche de nouvelles formes de transmission d'informations utiles à leur documentation aux gestionnaires par les usagers.

### ***Les interactions entre usagers et gestionnaires médiatisées par les collections numérisées***

*La consultation des collections en ligne peut être mise au service de la documentation des images s'il est possible d'intégrer des données produites par les usagers. C'est l'une des attentes du MNHN pour l'interface de consultation des collections de l'infrastructure e-ReColNat. Elle correspond à la transposition en ligne de l'activité de redétermination et de correction à laquelle se livrent les chercheurs travaillant dans les collections physiques. La possibilité existe déjà sur le site *JSTOR Global Plants* où sont laissés régulièrement par un petit noyau de contributeurs (seules cinquante personnes ont laissé plus de dix commentaires en trois ans) des précisions et des corrections sur les données associées aux images (ce en quoi consistent la majorité des commentaires laissés). Dans le panel de chercheurs et conservateurs interrogés, seule une personne, très investie dans la valorisation des herbiers français, a commenté régulièrement les images du site *JSTOR Global Plants* qui offre cette possibilité. La contribution à la correction ou à la précision des données du site est freinée du fait qu'aucune intégration automatique des apports des usagers n'est prévue.*

Les modalités d'ouverture de la base aux contributions des internautes ne sont pas consensuelles parmi les acteurs de l'infrastructure e-ReColNat. Il s'agit d'abord de savoir si l'image peut être considérée comme une « preuve » suffisante pour justifier la correction de la détermination d'un spécimen enregistrée dans la base. La question suscite d'autant plus de désaccords qu'elle est liée à celle de l'indépendance du document physique par rapport à la base de données. La planche d'herbier a jusqu'à présent tiré sa valeur de la possibilité de vérifier les déterminations faites à son égard, toutes consignées dans les étiquettes attachées au spécimen. Certains chercheurs interrogés, qui continuent d'accéder aux collections sans nécessairement consulter la base de données, signalent que dans le cas où ils se verraient obligés de passer par la base pour chercher les déterminations faites à propos de la planche par d'autres usagers, le document leur paraîtrait incomplet. On peut interpréter cette crainte comme une tension générée par la constitution de nouveaux « centres » dans l'environnement informationnel botanique, susceptible de concurrencer les collections physiques en offrant un accès facilité à une information enrichie des contributions des usagers. Certains conservateurs soulignent qu'il est toutefois indispensable de permettre l'enrichissement des données de la base en ligne, au profit de la valorisation des collections physiques, pour profiter des corrections apportées par la fréquentation des collections numérisées par des spécialistes.

## Conclusion

Les herbiers numérisés offrent un nouveau moyen d'accéder à une partie de l'information dont le spécimen est porteur - celle qui découle de la lecture des étiquettes des planches et de l'examen de ses caractères morphologiques observables sans analyse microscopique ni opération conduisant à en dégrader une partie. Les bases de données d'images, outre qu'elles comportent parfois des informations supplémentaires, comme les coordonnées de géolocalisation ou les publications mentionnant les spécimens, offrent de nouvelles possibilités pour la communication et l'exploitation des documents des collections ; la facilité de recherche de spécimens et de traitement des données qui y sont associées dépend toutefois de la finesse de l'indexation des images. L'herbier numérisé fournit un nouvel artefact médiateur de la communication scientifique, permettant la conservation des expertises et leur réappropriation par d'autres usagers éloignés dans le temps et l'espace, sans pour autant reproduire les fonctions de capitalisation des examens dont la planche est l'objet par sa documentarisation : la reproduction des fonctions du spécimen dans l'image trouve ses limites quand des modifications du document sont nécessaires pour en faire ressortir certaines propriétés ou y inscrire de nouvelles expertises. La structure transactionnelle de la communication scientifique, qui se traduit par la transformation des documents à mesure des nouvelles expertises, trouve son prolongement dans l'enrichissement des données associées aux images. L'accès facilité à l'information des collections profite en premier lieu à la documentation des images qui vise à améliorer l'efficacité de leur recherche. L'apport de la numérisation au pouvoir des centres documentaires d'amplifier l'information des spécimens tient autant à l'accès facilité aux informations par l'intermédiaire des images qu'à l'intégration des données produites par les usagers et le personnel des collections, essentielles à améliorer la documentation des images et à rendre efficace leur recherche.



## Références bibliographiques

Briet, Suzanne (1951), *Qu'est-ce que la documentation ?*, Paris : EDIT.

Chaudiron, Stéphane ; Ihadjadene, Madjid (2010), « De la recherche de l'information aux pratiques informationnelles », *Études de communication*, n°35, p. 13-30.

Després-Lonnet, Marie (2000), *Contribution à la conception d'interfaces de consultation de bases de données iconographiques*, Thèse en sciences de l'information et de la communication, Université de Lille 3 [en ligne] <https://tel.archives-ouvertes.fr/tel-01155546v1>

Després-Lonnet, Marie (2014), *Temps et lieux de la documentation. Transformation des contextes interprétatifs à l'ère d'internet*, Mémoire pour l'habilitation à diriger des recherches, Université de Lille 3 [en ligne] <https://tel.archives-ouvertes.fr/tel-01094789/document>

E-ReColNat (2013), ReColNat [en ligne], Infrastructure Nationale en Biologie et Santé e-ReColNat, date de consultation : janvier 2015, <http://ReColNat.org/>

Funk, Vicky (2003), « 100 uses for a herbarium: Well, at least 72 », *American Society of Plant Taxonomists Newsletter*, n°17, p. 17-19.

Heaton, Lorna ; Proulx, Serge (2012), « La construction locale d'une base transnationale de données en botanique », *Revue d'anthropologie des connaissances*, vol. 6, n° 1, p. 141-162.

Heaton, Lorna ; Millerand, Florence (2013), « La mise en base de données de matériaux de recherche en botanique et en écologie », *Revue d'anthropologie des connaissances*, vol. 7, n° 4, p. 885-913.

Hine, Christine (2008), *Systematics as Cyberscience*, Cambridge (Massachusetts), London (England) : The MIT Press.

Hine, Christine (2013), « The emergent qualities of digital specimen images in biology », *Information, Communication & Society*, vol. 16, n°7, p.1157-1175.

Latour, Bruno (1996), « Ces réseaux que la raison ignore - laboratoires, bibliothèques, collections », in Jacob, Christian ; Baratin, Marc (dir.), *Le pouvoir des bibliothèques. La mémoire des livres dans la culture occidentale*, Paris : Albin Michel, p. 23-46.

Jeanneret, Yves (2009), « Faire avec le faire communicationnel », *Actes Sémiotiques* [en ligne] <http://epublications.unilim.fr/revues/as/3336>, consulté en juin 2015

JSTOR (2000), JSTOR Global Plants [en ligne], Ithaka, date de modification : 2015, date de consultation : 2015, <https://plants.JSTOR.org/>

MNHN (2013), MNHN - Search [en ligne], MNHN, date de consultation : juin 2015, <https://science.mnhn.fr/institution/mnhn/search>

Nentwich, Michael (2003), *Cyberscience: research in the age of the Internet*, Vienna: Austrian Academy of Sciences Press.

Pérez, Eva ; Pignal, Marc (2013), « Numériser et promouvoir les collections d'histoire naturelle », *Bulletin des bibliothèques de France*, n°5, p.27-31.

The New York Botanical Garden (2007), Index Herbariorum: A Global Directory of Public Herbaria and Associated Staff [en ligne], The New York Botanical Garden, date de consultation : 2015, <http://sciweb.nybg.org/science2/IndexHerbariorum.asp>

Vinck, Dominique ; Natale, Enrico (2015), « La transformation des sciences historiques. La part du numérique », in Gorga, Adriana ; Leresche, Jean-Philippe (éd.), Transformations des disciplines académiques : entre innovation et résistance, Paris : EAC, p. 65-79.

Zacklad, Manuel (2004), « Processus de documentarisation dans les Documents pour l'Action (DopA) : statut des annotations et technologies de la coopération associées », Actes du colloque « Le numérique : Impact sur le cycle de vie du document pour une analyse interdisciplinaire », Montréal : Éditions de l'ENSSIB, [en ligne] [http://archivesic.ccsd.cnrs.fr/sic\\_00001072/](http://archivesic.ccsd.cnrs.fr/sic_00001072/), consulté en juin 2015

Zacklad, Manuel (2015), « Genres de dispositifs de médiation numérique et régimes de documentalité », in Gagnon-Arguin, Louise ; Maurel, Dominique ; Mas, Sabine (dir.), *Les genres de documents dans les organisations, analyse théorique et pratique*, Montréal : PUQ, p.145-189.

## Remerciements

Cette recherche a bénéficié du financement de l'Infrastructure Nationale en Biologie et Santé e-ReColNat (ANR-11-INBS-0004). Nous remercions les usagers, gestionnaires et conservateurs d'herbiers qui ont accepté de nous parler de leurs travaux et de leurs collections avec enthousiasme et patience, en particulier mais sans exhaustivité aucune, au MNHN, à l'AMAP et à l'herbier de l'université de Montpellier, ainsi qu'au Jardin Botanique et à la Société Linnéenne de Lyon.