



Compte-Rendu Scientifique INBS

INFORMATIONS SUR LE PROJET

Acronyme du projet

E-RECOLNAT

Rapport couvrant la période du

01/01/2020

au

31/12/2020

Date de notification de la convention attributive d'aide

2013-06-06

Titre complet du projet

Valorisation de 350 ans de collections d'histoire naturelle : une plateforme numérique

Mots clés

Collections naturalistes, Biodiversité, Taxinomie, Ecologie, Numérisation, Collections virtuelles, Accès, Environnement, Données ouvertes, Science ouverte, Universités, Muséums

Etablissement coordinateur

Museum National D'Histoire Naturelle

Date de début du projet

2013-01-01

Date de fin du projet

2022-12-31

Site web du projet

<https://www.recolnat.org>

**RESPONSABLE SCIENTIFIQUE ET TECHNIQUE DU PROJET,
REDACTEUR DU PRESENT RAPPORT**

Nom, Prénom

PIGNAL

Téléphone :

01 71 21 47 12

Courriel

pignal@mnhn.fr

Date de rédaction

2021-03-22

RESUME PUBLIC / PUBLIC SUMMARY

Résumé public et diffusable. Cette partie intègre les objectifs, l'avancement et les

faits marquants depuis le début du projet. Cette partie sera actualisée chaque année. Maximum 20 lignes - 2000 caractères.

L'infrastructure e-ReColNat produit, rassemble et rend accessible les données des collections naturalistes pour la communauté scientifique à travers plusieurs offres de services. Dans une logique de science ouverte, ces données sont exportables et réutilisables pour permettre de répondre aux grandes questions de recherche sur l'environnement et les changements globaux, la santé ou encore la sécurité alimentaire. Le périmètre de RECOLNAT englobe toutes les collections d'histoire naturelle conservées en France métropolitaine et les Outre-mer. Les 350 ans de collectes de spécimens issus de la géodiversité et de la biodiversité apportent une profondeur historique absolument unique pour exploiter ces sources de données. Le PIA e-ReColNat permet d'accélérer la production de données numériques et d'images : le corpus propose aujourd'hui plus de 10,8 millions de spécimens en accès libre. Ce corpus est aussi enrichi par le programme de sciences participatives des Herbonautes qui comprend plus de 5 millions de contributions depuis 2013. Par ailleurs, l'outil Annotate, à l'instar d'une paillasse virtuelle, permet d'effectuer des mesures morphométriques et des annotations sur les doubles numériques des collections. Enfin, le portail e-ReColNat est devenu le lieu de dépôt officiel des collections naturalistes des musées de France.

Depuis juin 2020, e-ReColNat a consolidé sa fondation sous la forme d'un groupement d'intérêt scientifique (GIS). L'ambition est de constituer un réseau d'excellence scientifique en France en rassemblant les professionnels et en harmonisant les pratiques liées aux collections naturalistes. Fin 2020, 18 membres et 2 partenaires parmi les muséums, universités et institutions de recherche avaient adhéré au GIS. RECOLNAT figure sur la feuille de route des infrastructures de recherche du MESRI depuis 2016 et constitue le point nodal français de l'infrastructure miroir européenne DiSSCo (Distributed Systems of Scientific Collections).

ETAT D'AVANCEMENT DU PROJET / PROGRESS OF THE PROJECT

Mise en place / Suivi (Lancement, organisation et gouvernance du projet, pilotage et dispositifs de suivi, recrutement des personnels, modalités d'accès à l'infrastructure). Maximum une page et demie, 6 300 caractères

il s'agit de la partie majeure du compte-rendu scientifique annuel. Décrire ici l'état d'avancement du projet par rapport à l'objectif initial. Les parties suivantes sont destinées à décrire l'organisation, avancées dans la réalisation du projet, formation, innovation, point sur les partenariats. Difficultés rencontrées. Mesures collectives. Perspectives pour l'année à venir.

WP1 : Gouvernance

Le WP1, composé par une équipe de 3 personnes de l'établissement coordinateur, assure le fonctionnement de la gouvernance avec l'organisation de ces comités, coordonne les travaux des autres WP et sert d'interlocuteur entre les comités de l'infrastructure, les tutelles, les financeurs et les administrations. Il assure la consolidation et rendu des rapports scientifiques, techniques et financiers aux différentes instances.

Principales actions et activités WP1

GIS RECOLNAT

En 2020 un livrable clé du PIA e-ReColNat et structurant pour la poursuite du programme a été rendu.

Sous l'impulsion du MNHN, des universités et des muséums français ont fondé en juin 2020 le groupement d'intérêt scientifique (GIS) du réseau national des collections naturalistes (RECOLNAT). Piloté par le Muséum national d'histoire naturelle, le GIS s'inscrit dans la continuité d'e-ReColNat et réunit 18 institutions. Son ambition est de constituer un réseau d'excellence scientifique dans le champ des collections naturalistes en rassemblant et mobilisant les professionnels, en développant des services d'accès aux données des spécimens, et en favorisant l'accès aux collections, tant physiques que virtuelles.

Une journée de lancement a été organisée le 4 novembre 2020. Compte tenu de la situation sanitaire, elle s'est déroulée par visioconférence, plus de 100 participants ont suivi les présentations et les échanges. En fin d'année, une première proposition de plan de travail du GIS a été validée par la gouvernance de l'infrastructure.

Liste des membres du GIS RECOLNAT : Muséum national d'histoire naturelle, Université Clermont Auvergne, Université Claude-Bernard Lyon 1, Université de Rennes 1, Muséum d'histoire naturelle de Dijon, Université de Bourgogne, Université Aix-Marseille, Université de Montpellier, Centre de coopération internationale en recherche agronomique pour le développement, Muséum d'histoire naturelle de La Rochelle, Muséum Aquarium de Nancy, Muséum d'histoire naturelle de Marseille, Université de Lorraine, Université de Grenoble, Institut de recherche pour le développement, Muséum d'histoire naturelle de Lille, Muséum d'histoire naturelle de Toulouse, Muséum d'histoire naturelle de Troyes, Sorbonne Université.

Évaluation des candidatures des muséums aux subventions du ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation (MESRI)

Pour la sixième année consécutive le WP1 a été l'interlocuteur du MESRI et a organisé une séance de travail pour que les membres du CS RECOLNAT expertisent et classent les 21 dossiers de candidatures de muséums aux subventions qui seront allouées par le ministère en 2021.

Évaluation des candidatures de l'appel à projet du programme de sciences participatives

Suite à l'appel à projet 2020 du programme les Herbonautes, le WP1 a coordonné

l'évaluations de candidatures par une partie de membres du CS RECOLNAT. Dix projets ont été retenus (voir WP6).

Candidatures à l'Equipex +

Le WP1 s'est fortement impliqué dans la rédaction du dossier de candidature à l'appel à manifestation d'intérêt d'équipement structurant pour la recherche, Equipex +. Coordonné par le Muséum national, le projet de candidature de 10 institutions partenaires a été soumis sous le nom e-COL+. Des 135 de projets déposés, 50 ont été retenus pour être financés et 2 autres ont été classés en liste complémentaire, e-COL + se trouve parmi ces 2 derniers.

Réunion des comités 2020

Comité de direction (CODIR) : 27/03/2020 et 04/11/2020

Comité scientifique (CS) : 19/02/2020 et 15/10/2020

Comité Exécutif (COMEX) : 04/11/2020

Depuis le début du programme, 28 comités ont été organisé afin de suivre l'avancement, veiller à l'exécution des travaux et donner les orientations de l'infrastructure.

Le comité de pilotage de suivi de l'opération de numérisation industrielle des herbiers a fait l'objet de 33 réunions afin de veiller au respect du calendrier et de la qualité de la prestation.

Travaux dans le cadre de l'infrastructure européenne DiSSCo

Le programme DiSSCo (Distributed System of Scientific Collections) a été inscrit sur la feuille de route du forum stratégique des infrastructures de recherche (ESFRI) en septembre 2018. Cet ambitieux programme européen qui rassemble déjà 120 établissements de 21 pays différents, vise à valoriser les collections naturalistes pour la recherche en les rendant accessible (notamment via leur numérisation), à homogénéiser les pratiques liées à leur conservation et leur accès, et à partager de meilleures pratiques collectives. Le principe sous-jacent d'accès aux données de bio-et géodiversité poursuivi par DiSSCo est celui - désormais classique - intitulé FAIR pour Findable (facile à trouver), Accessible (accessible), Interoperable (interopérable), Reusable (réutilisable).

La proximité des objectifs avec l'infrastructure nationale RECOLNAT la positionne naturellement comme un acteur important de la dynamique DiSSCo. C'est notamment à ce titre que RECOLNAT constitue le point nodal français de cette infrastructure. À travers ce qui est appelé le NN (National Node = nœud national), RECOLNAT fait le lien entre les

acteurs de son réseau national et le niveau des acteurs européens. En 2020, une visioconférence mensuelle réunit tous les NN des 21 pays européens engagés. Une page web sur la partie française de DiSSCo et le rôle de RECOLNAT est disponible à l'adresse : <<https://www.dissco.eu/fr/>>

Par ailleurs, RECOLNAT est impliqué activement dans le programme DiSSCo Prepare, visant à préparer l'opérationnalité de DiSSCo à l'horizon 2025/26. L'équipe RECOLNAT est notamment impliquée dans l'exercice de structurer le livre de coûts de l'infrastructure, sa gouvernance, l'engagement des parties prenantes et l'interopérabilité des systèmes et données.

Résultats obtenus dans l'année. Maximum une page et demie - 6 300 caractères.

WP2 : Infrastructure de publication et base de données

Acquisition et stockage des données

Les opérations de numérisation industrielle et de routine se sont terminées en 2019 et l'acquisition en 2020 a donc été restreinte. Un total de 36 900 nouveaux spécimens complets provenant de 4 muséums et 6 universités a été intégré via l'interface de saisie.

Le partenaire de numérisation Picturae a terminé de livrer les disques durs contenant les images TIFF au MNHN. Le transfert des disques vers l'architecture de stockage a dû être arrêté en avril 2020 en raison de la crise sanitaire.

Augmentation et publication des données

Les données de récolte de 17 935 spécimens ont été complétées et intégrées dans la base. Fin 2020, environ 10,8 millions de spécimens étaient disponibles en ligne. Ces données sont consultables et exportables en format JSON, CSV ou Darwin Core via l'interface explore.

Label RECOLNAT

La grande diversité des solutions logicielles de gestion des collections employée par les musées se traduit par une hétérogénéité dans le traitement des données. Afin de faciliter les opérations d'interopérabilité entre ces logiciels et la base RECOLNAT, une solution de labellisation a été créée par l'infrastructure. Ce label vise à normaliser l'information scientifique et s'appuie sur des normes et standards internationaux. En 2020, SNBase v.6 a obtenu son label argent et rejoint les autres éditeurs labellisés : Emu, Alienor et Webmuseum. Aussi, plus de 80 % du parc des solutions logicielles est labellisé. Ceci répond aux demandes faites par le CS concernant les demandes des muséums auprès du MESRI.

Mise en place d'un audit du SI RECOLNAT

Suite à des dysfonctionnements constatés dans le flux de données entre les bases, certaines instabilités et le départ du chef informatique, un audit d'analyse de l'existant a été mené afin d'obtenir des solutions. La résolution des problèmes pourrait être une TMA (tierce maintenance applicative), à engager en 2021.

WP3 : Collections d'herbiers

L'augmentation du corpus a été réalisée à partir de campagnes de numérisations industrielle et artisanale entre 2015 et 2019. Plus de 50% de la subvention du PIA, c'est-à-dire 8,5 M€ ont été dédiés aux tâches accomplies par ce WP.

Numérisation industrielle et attachage

Le bilan de production général est de 1 766 574 planches d'herbiers attachées appartenant à 207 collections issues de 33 institutions et 3 553 026 numérisées appartenant à 340 collections différentes issues de 43 institutions ont été traitées.

Numérisation 2020

Le partenaire UCA a mis en place un dispositif de numérisation de routine. L'opération est menée par Bibliothèque numérique de l'UCA et les images et données sont disponibles en ligne (<https://bibliotheque-virtuelle.bu.uca.fr/collection-tree/browse?collection=24>). À la fin de la prestation, ces spécimens viendront alimenter RECOLNAT.

2406 parts d'herbier ont été informatisées par MPU, ainsi que deux fonds d'iconographie botanique (vélin de Node-Véran et champignons de Delile).

Exploitation des collections numérisées

Différents types de travaux utilisent les collections du corpus RECOLNAT (cf. liste de publications). Parmi ces travaux, le projet IHMAGES a bénéficié de crédits FEDER et a pour objectif d'étudier l'évolution de la biodiversité du Massif central par l'informatisation de 100 000 données historiques issues des herbiers. Ce projet est coordonné par UCA en partenariat avec le MNHN, l'Université de Limoges, et les musées d'Aurillac, Montbrison, le Puy, Mende, Moulins, et Autun.

WP4 : Collections de types et figurés de Zoologie et Paléontologie

L'année 2020 a vu la fin de campagnes de numérisation mise en place par l'infrastructure

pour les collections des types et figurés en Zoologie et Paléontologie. Le bilan des opérations de numérisations réalisées entre 2014 et début 2019 est le suivant : 327 603 spécimens ont été informatisés dont 162 339 numérisés, ce qui représente un taux de réalisation respectif de 144% et 71%.

WP5 : Outils pour les communautés

Outil Annotate

Annotate permet de produire des données sur les images des spécimens des collections naturalistes à travers la documentation par annotation et la prise de mesures. Les éléments produits peuvent être exportés en une table de données structurées et des outils élémentaires de statistiques automatisés sont également disponibles.

En 2020, le développement pour compléter ses fonctionnalités a continué avec plusieurs améliorations : (i) gestion de plusieurs projets de recherche différents, partage et exports (ii) export/import de modèles Xper3, (iii) import d'images optimisé au local et aux métadonnées de smartphones, (iv) export de l'ensemble des métadonnées des images, (v) organisation d'un corpus en séquence documentée, avec titres et cartels, (vi) export des collections à la norme IIIF pour exposition (serveur mis en place).

Un élargissement du périmètre d'*Annotate* vers les sciences humaines et sociales (humanités numériques) et le deep learning a été mené en 2020. Une équipe de la BnF a commencé un travail de création des données d'apprentissage pour la reconnaissance automatique d'écriture musicale.

WP6 - Programme de science participative et animation des communautés

Les Herbonautes

Depuis 2013, *les Herbonautes* continuent à enrichir les métadonnées et données descriptives des spécimens. Ils contribuent à la complétude de la base de données RECOLNAT à partir des millions d'images des plantes d'herbiers numérisés et leurs étiquettes.

Lors de l'année 2020, 16 nouvelles missions ont été proposées à la communauté des *Herbonautes*, dont 13 ont été finalisées. Les 3 restantes seront terminées courant 2021.

En 2020, une première mission de transcription des collections paléontologiques a été lancée sous le libellé « *les Paléonautes* » avec l'objectif d'informatiser la collection Alphonse Péron.

Par ailleurs, dans le cadre du partenariat du projet IHMAGES, 4 missions ont été mise en place en 2020 avec l'objectif d'utiliser les informations contenues dans les herbiers anciens à des fins de connaissance de l'évolution de la végétation et des milieux du Massif central. Ce projet coordonné par l'UCA a bénéficié des crédits FEDER et apportera un co-financement à l'infrastructure de 21K€.

FICHIERS ILLUSTRANT L'AVANCEMENT DU PROJET

Afin d'illustrer l'avancement du projet, 4 fichiers en format .jpg peuvent être joints. Charger les fichiers en cliquant sur le lien "[Envoyer des fichiers](#)" ci-dessous.

```
[{"size":"0", "name":"Figure1.JPG", "ext":"jpeg"} {"size":"0", "name":"Figure2.JPG", "ext":"jpeg"} {"size":"0", "name":"Figure3.JPG", "ext":"jpeg"} {"size":"0", "name":"Figure4.JPG", "ext":"jpeg"}]
```

IMPACT SOCIO-ECONOMIQUE / SOCIO-ECONOMIC IMPACT

[Retombées pour la communauté scientifique \(Visibilité, adéquation avec les besoins, service rendu à la communauté académique et industrielle, qualité du service rendu...\).](#) Maximum 1 page - 4200 caractères.

Publications scientifiques et rapports

81 travaux faisant référence à l'IR RECOLNAT ont été publiés en 2020 (cf. liste des 10 travaux majeurs dans le paragraphe dédié). Ce chiffre est en augmentation et correspond à 7 % de l'ensemble de toutes les publications utilisant les collections naturalistes en France.

Doctorat

17 thèses de doctorat ayant utilisé l'infrastructure ont été soutenues durant l'année 2020, dans les domaines de la paléontologie, taxinomie intégrative, mammalogie, phylogénétique, etc.

Intervention dans le Master Systématique, Évolution, Paléontologie (Sorbonne Université/MNHN)

L'outil d'annotation et de mesures sur les images d'herbiers va apporter une nouvelle façon de travailler aux systématiciens et, plus généralement, aux botanistes. Les collections virtuelles se montrent ici comme des compléments efficaces des collections physiques, sans les remplacer cependant. Elles permettent à la fois une facilité d'accès (particulièrement utile durant le confinement), d'augmenter la vitesse d'acquisition (près de 1 000 mesures par jour) et ne porte pas atteinte à la conservation des objets physiques. Afin de présenter cet outil aux étudiants, plusieurs sessions de formations ont eu lieu en 2020.

Publication du livre blanc sur les infrastructures du domaine environnement (Allenvi)

La publication du livre blanc sur les infrastructures de recherche du domaine des sciences du système Terre et de l'environnement a permis de donner une visibilité à l'IR RECOLNAT au sein du paysage stratégique national et européen.

Méta-inventaire des collections entomologiques des Pays de la Loire

Dans le cadre d'une opération pilote de recensement des collections d'insectes des Pays de la Loire financée en 2019 (crédits SOERE-AllEnvi), une présentation et une publication avait été prévue en 2020. Hélas, le confinement et la crise sanitaire n'ont pas permis la tenue du colloque. Ce travail est repoussé en 2021.

Pour rappel, cette mission a permis de recenser les contenus taxinomiques de plus de 10 000 boîtes conservées dans 15 institutions publiques, et plus de 5100 issues de 50 collections particulières. Cette mission s'inscrit dans la continuité du recensement des collections d'herbiers mené de 2014 à 2018 et préfigure une des missions du futur GIS RECOLNAT. La visibilité des collections est en effet le premier élément qui permet leur valorisation pour la recherche scientifique. Cette approche est complémentaire de celle pratiquée au niveau européen dans le cadre d'ICEDIG, programme visant à préparer l'infrastructure européenne DiSSCo et auquel RECOLNAT avait participé pour la partie française.

Expertise des collections naturalistes du Département des Hautes-Alpes

Une mission d'expertise a été commanditée auprès de l'infrastructure par le Département des Hautes-Alpes. Il s'agissait d'établir un diagnostic sur les collections naturalistes propriété du Département (bilan chiffré, état d'informatisation et de numérisation, état de conservation général, statut et adéquation avec les différents cadres réglementaires afférents, diversité et intérêt scientifique, technique et patrimonial), d'apporter une assistance technique afin de valoriser les collections naturalistes du Département des

Hautes-Alpes sur le portail RECOLNAT et de rédiger un rapport synthétique présentant le diagnostic et les résultats de l'intégration, ainsi qu'une vision prospective sur les suites à mener sur cet ensemble de collections.

Cette partie prospective doit servir à accompagner les choix politiques en matière d'acquisition, de mutualisation, de mise aux normes ou de numérisation/informatisation des collections d'une part, et d'accompagnement et de formations à mettre en place sur les personnels d'autre part. L'ensemble permettra de nourrir le projet scientifique et culturel du Musée muséum départemental en cours de rédaction par ses services.

Par ailleurs, une restitution de ce rapport auprès des administrés et des élus du Département des Hautes-Alpes a eu lieu en décembre 2020 afin d'aider aux prises de décision locales en matière de collections naturalistes (validation PSC, recrutements, budgets).

Formation professionnelle (Types de formation, quels opérateurs? quels utilisateurs?). Maximum 1 page - 4200 caractères.

Programme conjoint de formation avec l'Office de coopération et d'information muséales (Ocim)

Plusieurs réunions de travail ont été effectuées avec l'OCIM pour mener des actions conjointes de formation à destination des membres du réseau des collections naturalistes. Hélas, la crise sanitaire n'a pas permis de les rendre effective en 2020.

Dans le but de structurer le contenu de futures formations, RECOLNAT a participé aux enquêtes conduites par l'OCIM sur les muséums et sur les collections dans les universités et les grandes écoles. Par ailleurs, un premier recensement des thèmes de formation qui serait souhaitable au sein de l'infrastructure a été mené.

Formation aux dossiers de demande de Fonds européen de développement régional (FEDER)

La recherche de fonds afin de conserver et valoriser les collections d'histoire naturelle est une activité partagée par tous les acteurs. Les financements européens existent, mais la culture pour aller les chercher et la technicité pour monter correctement un dossier manquent à de nombreux partenaires. Avec l'expertise apportée par l'université Clermont-Auvergne qui a déjà remporté le financement de plusieurs programmes FEDER, nous avons eu plusieurs entretiens et réunion de concertation pour organiser une journée d'échanges et de formation sur ces dispositifs. Il était envisagé l'organisation d'une journée en fin 2020 mais elle a été reportée en raison de la crise sanitaire.

Journée du réseau botanique

Cette journée n'a pas pu se tenir en 2020 en raison de la crise sanitaire. Elle avait été prévue à Angers au mois de juin 2020.

[Partenariats avec des entreprises pour des développements technologiques et relations contractuelles établies \(Transferts de technologies, partenariats, créations de start-up\). Maximum 1 page - 4200 caractères.](#)

L'infrastructure RECOLNAT se situe très en amont des processus industriels puisqu'elle propose une offre de services dont le cœur est constitué par la mise à disposition des données descriptives et analytiques sur la biodiversité et la géodiversité issues de collections publiques à destination de la communauté scientifique. Depuis la loi du 7 octobre 2016, dite Loi Lemaire, le principe de l'open data est posé par défaut, ce qui permet un meilleur accès aux données publiques. Même si la Loi Valter permet aux établissements culturels de lever des redevances d'accès auprès des utilisateurs pour la réutilisation des données/métadonnées numérisées, cette démarche se révélerait contre-productive puisque les établissements membres sont de grands utilisateurs des ressources qu'ils y déposent.

L'infrastructure RECOLNAT s'intéresse fortement à son modèle économique. Elle s'investit notamment en position de leader au sein du groupe de travail sur cette thématique dans le programme DiSSCo Prepare, en préparation de l'infrastructure miroir DiSSCo. Le modèle économique des IR des collections naturalistes n'est pas fondé sur le corpus existant mais sur la production de nouvelles données pour la recherche (*digitisation on demand*).

Par ailleurs, la nouvelle feuille de route du GIS montre des attentes fortes liées aux services offerts par l'infrastructure dans les établissements du périmètre, notamment les collectivités territoriales. Une ambition émergente en 2020 est de diversifier et consolider l'offre de services et ainsi augmenter les dispositifs de prestations et d'accompagnements au niveau national.

Enfin, RECOLNAT participe au groupe de travail « Bonnes pratiques » piloté par le MESRI depuis 2018, ce qui offre un espace d'interactions et d'échanges autour des relations avec le monde industriel.

Échanges avec les éditeurs de bases de données muséologiques

Les muséums territoriaux utilisent des bases de données commerciales surtout fondées sur

des modèles du monde des musées de beaux-arts pour lesquels la documentation des objets de collection ne repose pas sur les mêmes fondements scientifiques. Ces bases ne suivent notamment pas de standards internationaux en matière de diversité biologique (nomenclature, taxinomie, index géographiques, matériaux, etc.).

Le MESRI ayant lié l'octroi de crédits de numérisation à la compatibilité de ces bases de données avec la base e-ReColNat, il a été défini deux labels de compatibilité. Le premier (argent) valide la capacité de verser des données vers la base nationale e-ReColNat et le second (or) celle de pouvoir échanger en interopérabilité des données entre e-ReColNat et les bases de données locales.

Les sociétés étant parfois réticentes à adapter leurs produits, des entretiens ont été organisés avec toutes les entreprises, avec l'aide de membres du Comité Scientifique de manière à expliquer l'importance des développements nécessaires. À la suite des entretiens, plusieurs de ces sociétés ont développé des adaptations. Début 2020, près de 70% des muséums territoriaux sont en capacité de transférer leurs données et images vers la base de données nationale RECOLNAT. Toutefois, la fin de la prestation de numérisation a entraîné beaucoup de difficultés informatiques qui ont empêché l'intégration de ces données.

LISTE DES 10 PUBLICATIONS ET TRAVAUX MAJEURS DU PROJET

[Renseigner la liste des 10 publications et travaux majeurs, selon vous, issus de l'utilisation de l'infrastructure et publiés dans une revue référencée dans le web of science au cours de l'année. Seules les publications effectives \(avec date de publicatio](#)

1. Bellis E.S., Kelly E.A., Lorts C.M., Gao H., DeLeo V.L., Rouhan G., Budden A., Bhaskara G.B., Hu Z., Muscarella R., Timko M.P., Nebie B., Runo S.M., Chilcoat N.D., Juenger T.E., Morris G.P., DePamphilis C.W. & Lasky J.R. 2020. — Genomics of sorghum local adaptation to a parasitic plant. *Proceedings of the National Academy of Sciences* 117 (8): 4243-4251. <https://doi.org/10.1073/pnas.1908707117> IF 9,351
2. Bolotov I.N., Konopleva E.S., Vikhrev I.V., Gofarov M.Yu., Lopes-Lima M., Bogan A.E., Lunn Z., Chan N., Win T., Aksenova O.V., Tomilova A.A., Tanmuangpak K., Tumpeesuwan S. & Kondakov A.V. 2020. — New freshwater mussel taxa discoveries clarify biogeographic division of Southeast Asia. *Scientific Reports* 10 (1): 6616. <https://doi.org/10.1038/s41598-020-63612-5> IF 3,998
3. Collareta A., Lambert O., Muizon C. de, Benites Palomino A.M., Urbina M. &

- Bianucci G. 2020. — A new physeteroid from the late Miocene of Peru expands the diversity of extinct dwarf and pygmy sperm whales (Cetacea: Odontoceti: Kogiidae). *Comptes Rendus Palevol* (5). <https://doi.org/10.5852/cr-palevol2020v19a5> IF 1,624
4. Condamine F.L., Nel A., Grandcolas P. & Legendre F. 2020. — Fossil and phylogenetic analyses reveal recurrent periods of diversification and extinction in dictyopteran insects. *Cladistics*: cla.12412. <https://doi.org/10.1111/cla.12412> IF 4,727
 5. Ewin T.A.M. & Gale A.S. 2020. — Asteroids (Echinodermata) from the Barremian (Lower Cretaceous) of Agadir Basin, west Morocco. *Journal of Paleontology* 94 (5): 931-954. <https://doi.org/10.1017/jpa.2020.20> IF 1,653
 6. Jabbour F., Du Pasquier P.-E., Chazalviel L., Le Guilloux M., Condee Silva N., Deveaux Y., Manicacci D., Galipot P., Heiss An.G. & Damerval C. 2020. — Evolution of the distribution area of the Mediterranean *Nigella damascena* and a likely multiple molecular origin of its perianth dimorphism. *Flora*. <https://doi.org/10.1016/j.flora.2020.151735> IF 1,591
 7. Jauvion C., Audo D., Bernard S., Vannier J., Daley A.C. & Charbonnier S. 2020. — A new polychelidan lobster preserved with its eggs in a 165 Ma nodule. *Scientific Reports* 10 (1): 3574. <https://doi.org/10.1038/s41598-020-60282-1> IF 3,998
 8. Jirapatrasilp P., Tongkerd P., Jeratthitikul E., Liew T.-S., Pholyotha A., Sutcharit C. & Panha S. 2020. — Molecular phylogeny of the limacoid snail family Dyakiidae in Southeast Asia, with the description of a new genus and species. *Zoological Journal of the Linnean Society*. <https://doi.org/10.1093/zoolinnean/zlaa129> IF 2,824
 9. PETZOLD A., MAGNANT A.-S., EDDERAI D., CHARDONNET B., RIGOULET J., SAINT-JALME M. & HASSANIN A. 2020. — First insights into past biodiversity of giraffes based on mitochondrial sequences from museum specimens. *European Journal of Taxonomy* 703: 1-33. <https://doi.org/10.5852/ejt.2020.703> IF 1,393
 10. Rouhan G. & Gaudeul M. 2020. — Plant Taxonomy: A Historical Perspective, Current Challenges, and Perspectives, in Besse P. (ed.), *Molecular Plant Taxonomy* Vol. 2222. New York, NY, Springer US. p. 1-38. (Methods in Molecular Biology). https://doi.org/10.1007/978-1-0716-0997-2_1 IF 12,084

SUITES DONNEES AUX RECOMMANDATIONS EMISES PAR LE JURY INTERNATIONAL

Il s'agit de décrire les actions et mesures prises suite aux recommandations émises par le jury dans le cadre de l'évaluation. Maximum 1 page - 4 200 caractères.

Lors de l'évaluation finale, le jury s'est déclaré inquiet quant aux critères de structuration scientifique de l'infrastructure et a demandé que deux points structurants soient développés : la promotion de l'infrastructure et sa consolidation.

Promotion de l'infrastructure

Structuration de la communauté au niveau français

Par ses activités de développement et d'animation de réseaux, l'infrastructure RECOLNAT poursuit la promotion des services qu'elle offre aux usagers des collections naturalistes. Une des ambitions est d'accroître cette promotion, afin de sensibiliser un panel plus important d'utilisateurs, et de permettre une valorisation optimisée des corpus et des services proposés. La constitution du GIS RECOLNAT en juin 2020 vient renforcer concrètement cette promotion et structure les communautés conservant et étudiant des collections naturalistes tout en accroissant leur visibilité. Par ailleurs, l'année 2020 montre un nombre d'utilisateurs actifs de l'infrastructure numérique en hausse (15391 utilisateurs actifs en 2020 contre 11169 en 2019). Une demi-journée de promotion de l'infrastructure RECOLNAT, et de bilan du programme e-ReColNat, a eu lieu le 4 novembre 2020. Plus de 100 personnes d'institutions différentes y ont participé.

Visibilité européenne

RECOLNAT est assuré d'une visibilité européenne en étant le point nodal français de l'infrastructure DiSSCo. L'IR a été sollicitée à plusieurs reprises pour exposer sa structuration aux autres partenaires. De même, Dimitris Koureas, directeur exécutif de DiSSCo, s'est adressé aux nœuds nationaux d'Europe (30 janvier 2020) en disant "that the next time he is asked what DiSSCo wants to achieve he can point to the French activities and their level of scoping and organisation. He was especially impressed with the level of services which could be powerful tools and software to be embedded under the umbrella of the DiSSCo data model and could thus be opened to global user communities".

Alliance avec des spécialistes du machine learning et de l'analyse d'images

En 2020, l'IR a travaillé durant le premier semestre à la réponse et à la constitution du dossier de l'Equipex+. Le projet comprend le développement d'outils de deep learning et de machine learning pour valoriser le corpus existant associé au financement de la conservation pérenne des images et des données, ainsi que l'enrichissement du corpus disponible en imagerie 3D.

Le projet e-COL+ a été classé A et fut sélectionné parmi les bénéficiaires, mais sans budget dédié à ce stade.

Meilleure ergonomie du site web pour les spécialistes et les non-spécialistes

La priorité donnée à l'intégration des images issues de la numérisation de masse d'une part, et la fondation du GIS RECOLNAT d'autre part, n'ont pas permis de revoir l'ergonomie

du portail web d'accès aux collections. L'audit sur l'architecture et les flux de données de RECOLNAT a été réalisé dans le courant de 2020. Une consolidation et des prestations de services (TMA) doivent être mises en œuvre en 2021.

Consolidation de l'infrastructure

Les efforts ont porté en 2019 et début 2020 sur le regroupement du corpus d'images produits dans le cadre de l'infrastructure. L'inventaire des données et des images produites par d'autres acteurs est un axe qui a été en partie réalisée en 2020 (hélas impacté par la crise sanitaire). Les outils d'interopérabilité, de consultation et de saisie des données ont fait l'objet d'un audit en 2020 pour une mise en œuvre en 2021/22.

Plan de conservation pérenne des données

Le plan de conservation des données a été initié par la définition avec le CINES des conditions de transfert et de conservation comme indiqué dans le paragraphe décrivant les activités du WP2. Sa mise en œuvre a commencé par le transfert des images depuis des disques durs amovibles vers un serveur local au MNHN offrant de meilleures conditions de conservation.

COMMENTAIRES LIBRES / FREE COMMENTS

[Ces commentaires libres peuvent porter sur le projet lui-même et sa trajectoire, sur les indicateurs fournis, sur les aspects financiers... Maximum 2 pages - 8 400 caractères.](#)

1/ Le nombre de projets traités est difficile à produire dans le cadre d'une infrastructure numérique. Nous avons indiqué ici le nombre de projets portés par l'infrastructure (missions Herbonautes : 16, expertise des collections dans les Hautes-Alpes : 1, méta-inventaire des collections entomologiques des Pays de la Loire : 1).

2/ Le nombre de requêtes traitées est fondé sur la fréquentation par des utilisateurs actifs sur le site d'accès à l'infrastructure (portail Explore) ; il est nette progression en 2020 (+ 37 % depuis 2019, sans doute en lien avec la COVID-19).

3/ Le cofinancement de 16000 € noté ci-dessous dans les indicateurs 2020 comme émanant du MNHN correspond aux cotisations des membres du GIS RECOLNAT (porté administrativement par MNHN).

RECOLNAT est organisé pour demeurer le point nodal français de l'infrastructure miroir européenne DiSSCo. Cet alignement stratégique passe par une structuration de la communauté métier et l'optimisation des services offerts aux utilisateurs. Ainsi, seul un effort coordonné des institutions détentrices d'objets de l'histoire naturelle permettra la transition vers une gestion des données harmonisée par la définition de

pratiques et de standards communs. Bien que les collections physiques constituent des sources indispensables et irremplaçables, les outils permettant l'étude virtuelle des objets et l'exploitation en masse de leurs données ouvrent des perspectives nouvelles. L'émergence de la notion de spécimen digital désigne la connexion et l'interopérabilité de toutes les informations relatives à un spécimen (morphologie, génétique, etc.). Ce nouvel algorithme de regroupement donne un avant-goût des possibilités d'intégration complète des données de biodiversité d'un futur proche. L'intégration d'images 3D de modèles biologiques compte également parmi ces innovations. L'érosion accélérée de la biodiversité renforce le rôle crucial de l'exploitation et de la documentation des données des collections.

Les perspectives 2021 de l'infrastructure sont les suivantes :

- Acquisition de nouvelles données et consolidation de l'architecture informatique et de la gestion des flux de données avec les muséums et universités
- Conservation pérenne des données et images HD
- Veille sur les nouvelles technologies (spécimen digital, intégration images 3D, ELViS, IA)
- Renforcement des interactions avec d'autres infrastructures de données sur la biodiversité (PNDB, RARe et iBOL).

INDICATEURS 2020

Taux de réalisation depuis le début du projet(%) :

(toutes les données sont renseignées en année civile, sauf mention contraire)

92

Formations, Diffusion:

Nombre de personnes formées aux techniques de la plateforme*	250
Nombre de fois ou l'infrastructure a été présentée lors de congrès ou lors de salons**	2
Monographies, ouvrages collectifs ou actes mentionnant le soutien financier du PIA pour le projet	0

Publications:

Publications mentionnant le soutien financier du PIA pour le projet, résultant de l'utilisation de la plateforme*	81
Publications mentionnant le soutien financier du PIA pour le projet, liées à la conception de la plateforme**	0

Utilisation de l'infrastructure (Projets):

L'infrastructure peut choisir de ne compléter qu'un des deux items (projets/requêtes)

Nombre de projets traités*	18
Volume horaire moyen d'un projet (en jours)	40

Ecart type	1
------------	---

Utilisation de l'infrastructure (Requêtes):

L'infrastructure peut choisir de ne compléter qu'un des deux items (projets/requêtes)

Nombre de requêtes traitées *	15391
Nombre de demandes *	15391

Utilisation de l'infrastructure (Taux d'utilisation):

En fonction des INBS, remplir un ou deux des tableaux (Taux d'utilisation/Taux de pénétrance). Conserver le même périmètre d'une année à l'autre.

Temps d'occupation machine *	365
Ecart type	VIDE
Délai d'attente maximum entre la réception de la demande et la livraison (en jours)	VIDE

Taux de pénétrance :

En fonction des INBS, remplir un ou deux des tableaux (Taux d'utilisation/Taux de pénétrance). Conserver le même périmètre d'une année à l'autre.

Taux de pénétrance par rapport à une cible définie**	VIDE
--	------

Pérennisation:

Nombre d'ETPT nécessaires afin de pérenniser l'infrastructure*	5
Nombre d'ETPT stabilisés au cours de l'année 2020	0
Nombre d'ETPT stabilisés du début du projet au 31/12/2020**	2

Brevets :

Nombre de brevets déposés dans l'année	0
--	---

Valorisation et ressources propres:

Montant des ressources engendrées par la PI (en k€)*	0
--	---

Montant des ressources propres hors PI (en k€)**	447700
Montant des recettes générées au cours de l'année à la suite de l'utilisation de l'infrastructure par des tiers extérieurs académiques (en k€)***	0
Montant des recettes générées au cours de l'année à la suite de l'utilisation de l'infrastructure par des tiers extérieurs non académiques (en k€)***	0

COFINANCEMENTS PAR ETABLISSEMENT PARTENAIRE

Liste des cofinancements :

Etablissement Coordinateur / Partenaire ayant obtenu le financement	Type cofinancier *	Nom cofinancier	Nature / Objet du financement **	Montant sur lequel le cofinancier s'est engagé sur la durée du projet (en €)	Montants perçus en 2020 (en €)
Museum National D'Histoire Naturelle	PUBLIC-ADMINISTRATION	ALLENVI	monétaire	35000	35000
Museum National D'Histoire Naturelle	PUBLIC-ADMINISTRATION	COMUE USPC (Université Sorbonne Paris Cité)	nature	10000	10000
Museum National D'Histoire Naturelle	PUBLIC-ADMINISTRATION	Muséum national d'histoire naturelle (MNHN)	monétaire	16000	16000
Université de Dijon Bourgogne	PUBLIC-COLLECTIVITES	Région Bourgogne-Franche-Comté	monétaire	161000	89555
Totaux				222000	150555

DONNEES FINANCIERES LIEES A LA SCIENCE OUVERTE - INDICATEURS

Montant des APC (Articles Processing Charges) payés dans le cadre du projet

0

Coût associés à la gestion des données du projet (stockage, gestion, mise à disposition,...)

265775

DONNEES FINANCIERES LIEES A LA SCIENCE OUVERTE - FICHER

* Indiquez ici le detail du montant global. Compléter et joindre le fichier Excel : Cliquer [ICI](#). Une fois les informations renseignées charger le fichier (au format .xlsx) en cliquant sur le lien "Envoyer des fichiers" ci-dessous.

```
[{"size":"0", "name":"ScienceOuverte_RECOLNAT_2020.xlsx", "ext":"xlsx"}]
```

VALIDATION

[En cochant cette case je certifie avoir complété l'ensemble des informations concernant le projet](#)

oui

Principaux graphiques illustrant le travail accompli par e-ReColNat

Organigramme fonctionnel du PIA e-ReColNat.

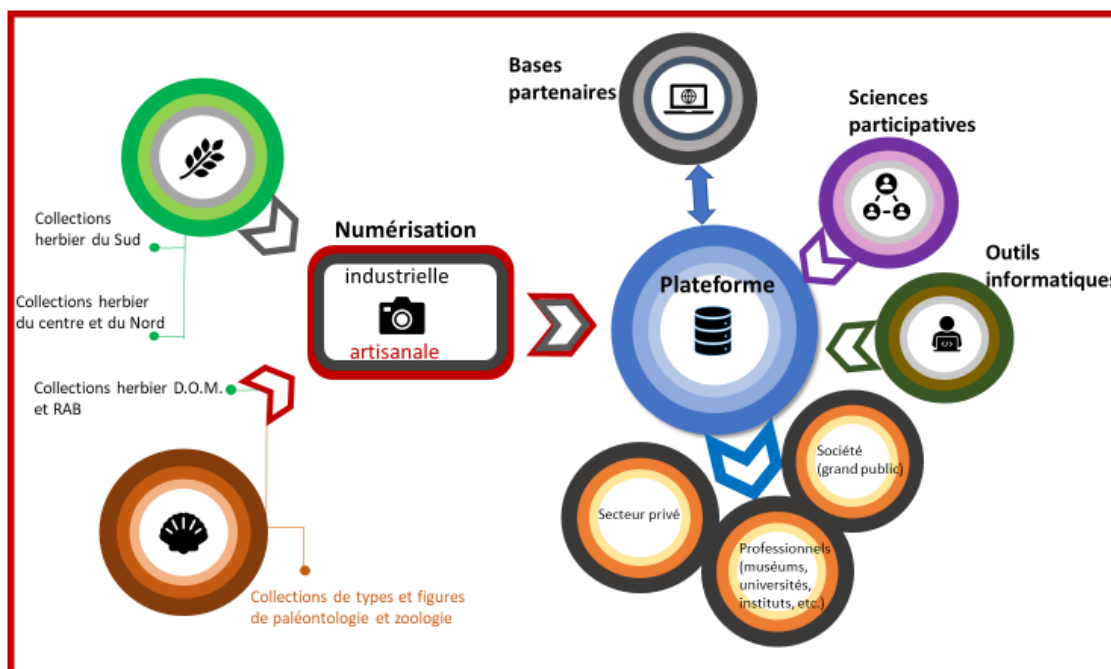
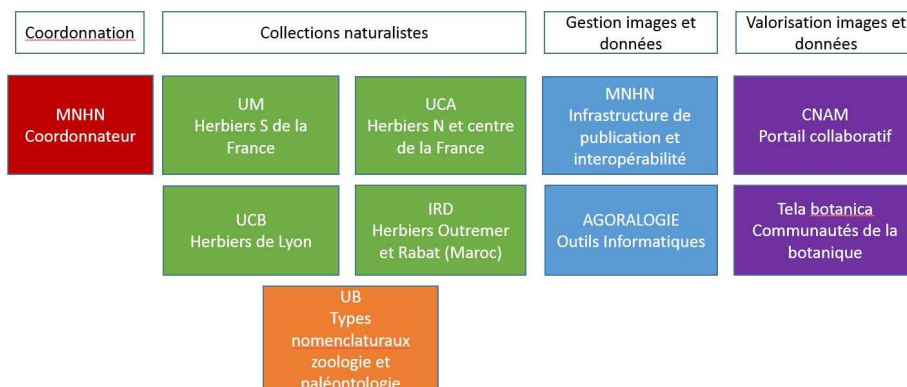


Schéma des composantes de l'infrastructure et les champs d'intervention des WPs.

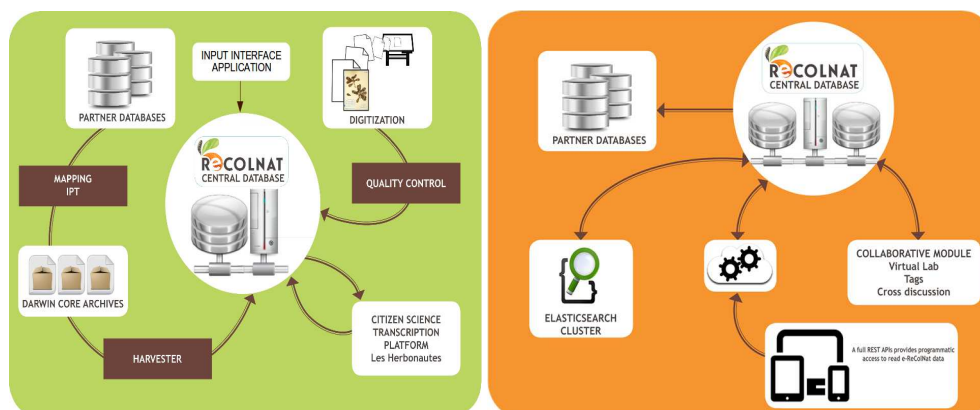
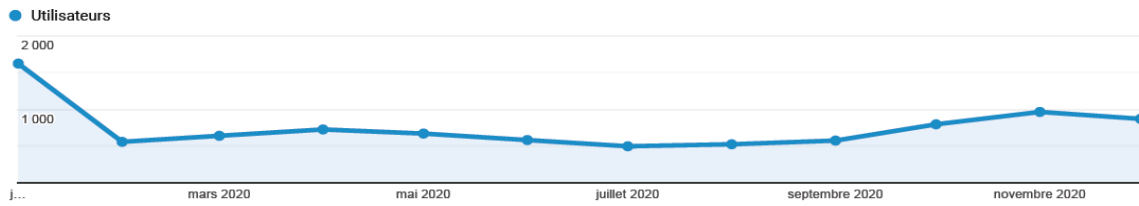


Schéma de la circulation de la donnée : acquisition, stockage, augmentation et publication

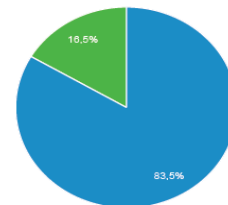
Tous les utilisateurs
100,00 %, Utilisateurs

1 janv. 2020 - 31 déc. 2020

Vue d'ensemble

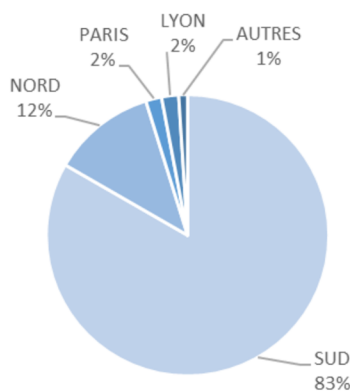


New Visitor Returning Visitor

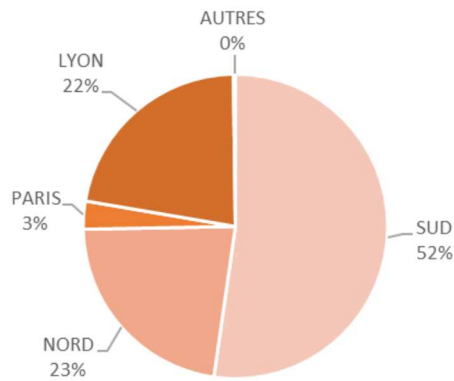


Statistiques 2020 de l'interface de consultation des collections Explore (<https://explore.recolnat.org>)

ATTACHAGE



NUMERISATION

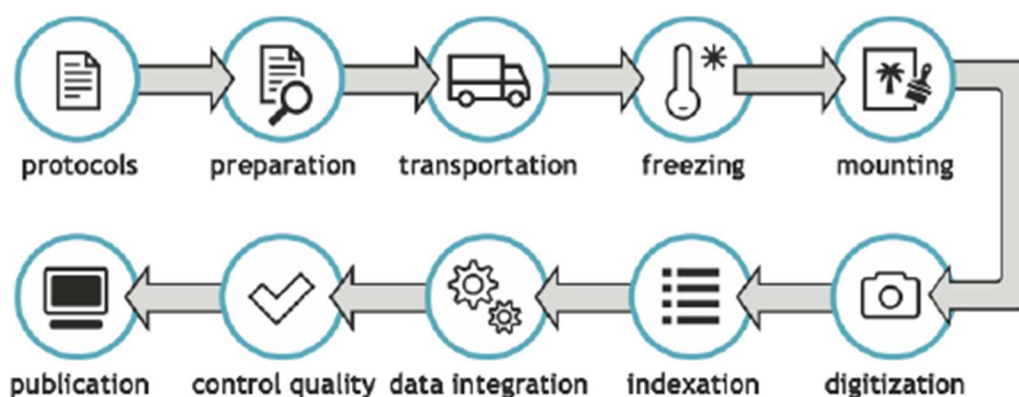


Pourcentage et répartition de la production d'attachage et numérisation d'herbiers.

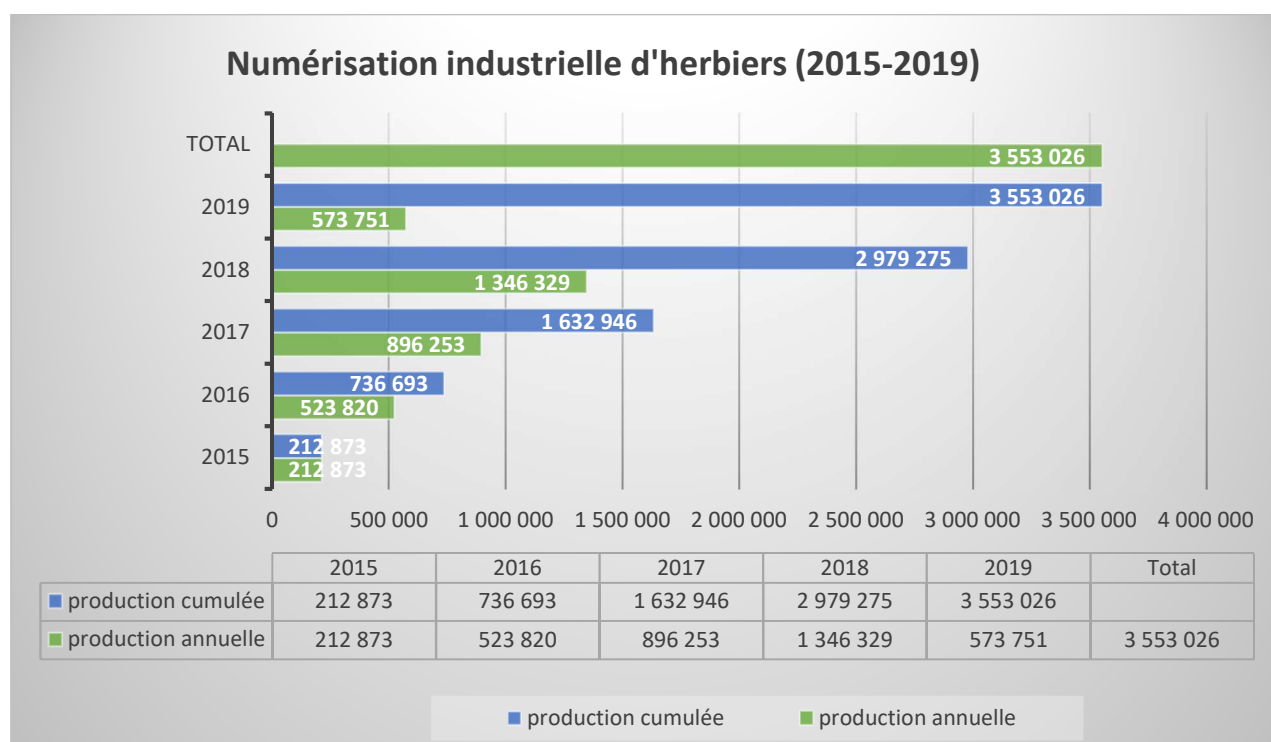
	2015	2016	2016	2016	2016	2017	2017	2017	2017	2018	2018	2018	2018	2019	2019
SITE	TRIMESTRE 4	TRIMESTRE 1	TRIMESTRE 2	TRIMESTRE 3	TRIMESTRE 4	TRIMESTRE 1	TRIMESTRE 2	TRIMESTRE 3	TRIMESTRE 4	TRIMESTRE 1	TRIMESTRE 2	TRIMESTRE 3	TRIMESTRE 4	TRIMESTRE 1	TRIMESTRE 2
SJDV															
PARIS															
AUBIERE															
LYON															
HEILOO															

ATTACHAGE NUMERISATION ATTACHAGE +|NUMERISATION

Sites de production de l'opération de traitement d'herbiers



Cycle de la chaîne opératoire du traitement des collections d'herbiers

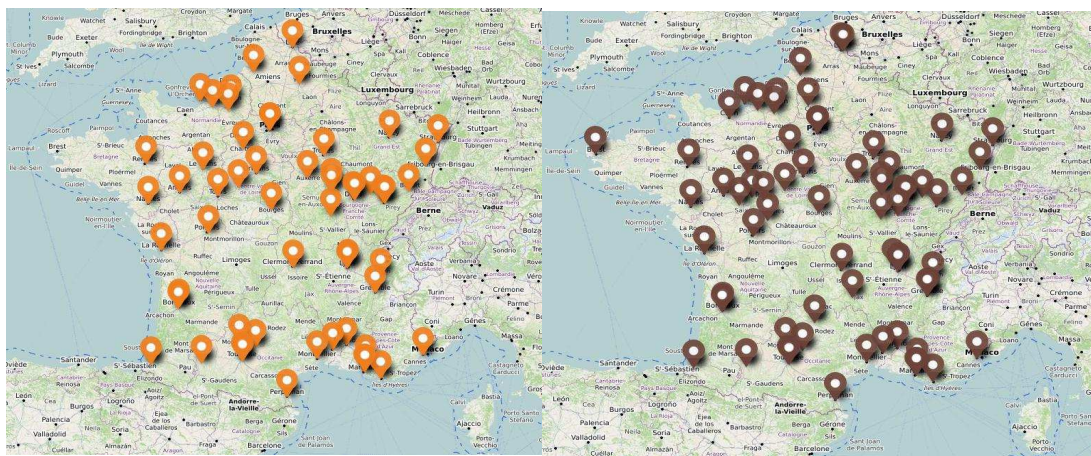


Production totale de numérisation industrielle d'herbiers réalisée par l'infrastructure.

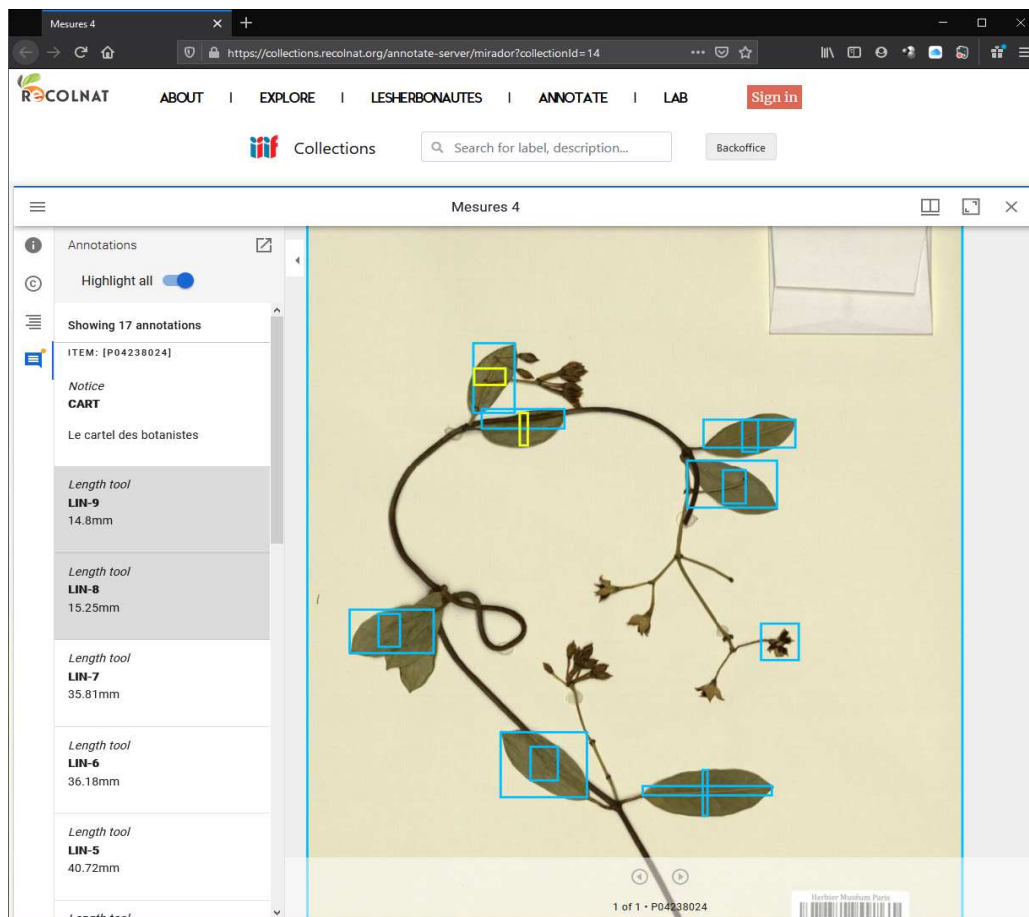
Numérisation de routine DOM et Maroc	
Nom des herbiers partenaires	Nombre de numérisations
Guyane (CAY)	74 500
Nouvelle Calédonie (NOU)	64 500
Rabat, Maroc (RAB)	41 000
Réunion (REU)	23 000
Guadeloupe (GUAD)	5 000
Saint Pierre et Miquelon	1 600
Total de spécimens numérisés	209 600



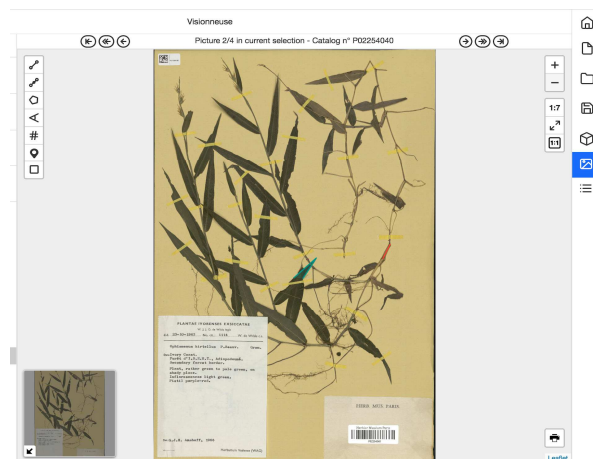
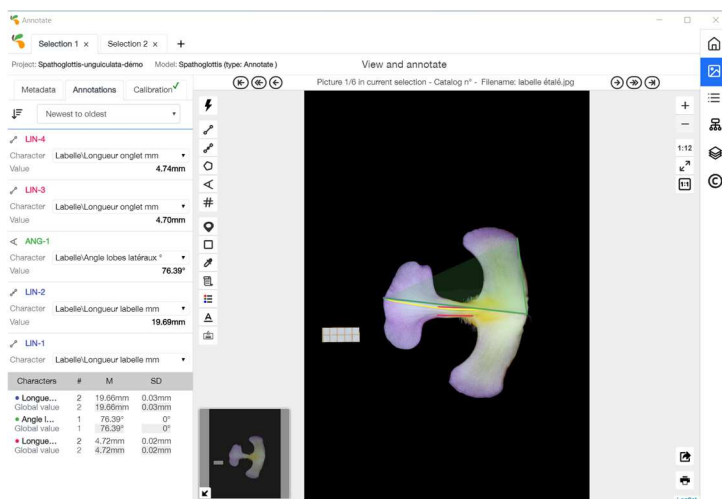
Collections de l'herbier de Clermont-Ferrand, projet IHMAGES



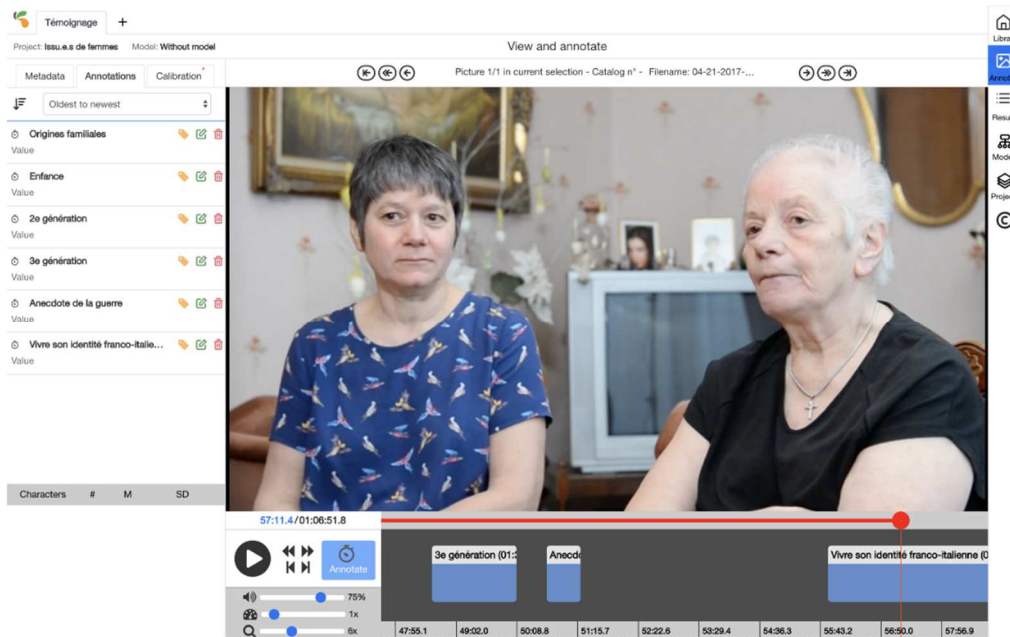
Établissements conservant des types et figurés zoologiques (orange) et paléontologiques (marron).



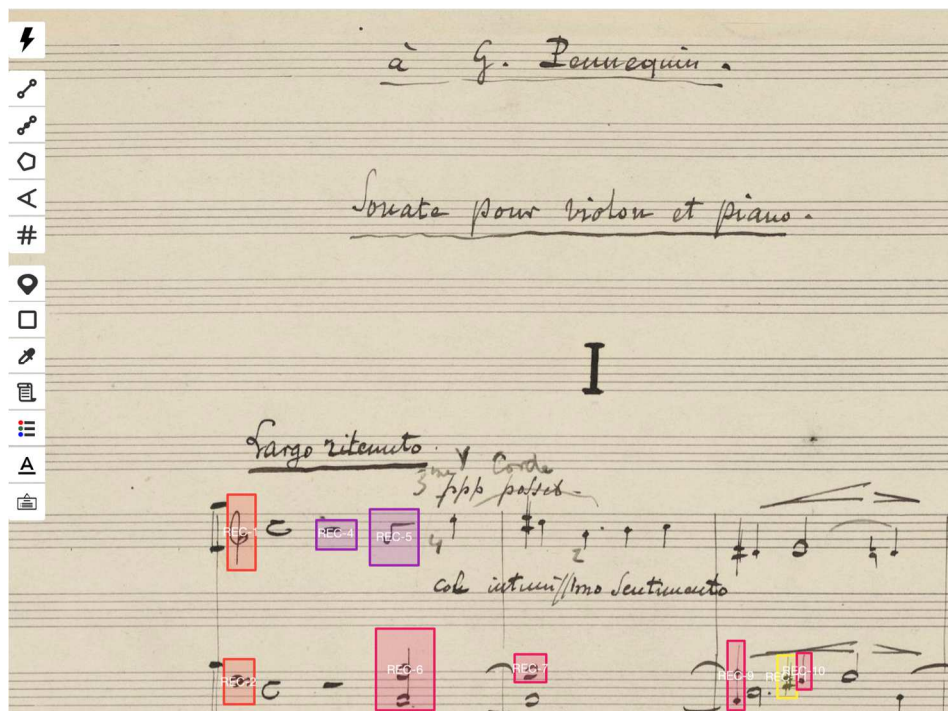
Exposition générée par Annotate en IIF d'images avec leurs annotations.



Prise de mesures, annotations et « tagage » à partir des bibliothèques d'images et/ou des photos prises sur le terrain.



Copie d'écran d'utilisation de la partie annotations de vidéo. Ici un entretien dans le cadre du projet « Issu.e.s de femmes » autour de l'immigration.



Annotations en écriture musicale effectuées avec Annotate pour des données d'apprentissage en deep-learning appliquées à de la reconnaissance automatique dans des partitions manuscrites (base BnF).

Les herbonautes
L'herbier numérique collaboratif citoyen

Qui sommes-nous - Missions - Discussions - Accès rapide

Etape 2 dernière - 200 ans de collections botaniques dans le Massif central

Mission terminée
Une dernière mission pour cette étape !

Nombre de spécimens	7571
Contributions	15240
Chef de mission	Vero
Ouverture	07 juillet 2020

Mission terminée

7
membres

7571
spécimens vus

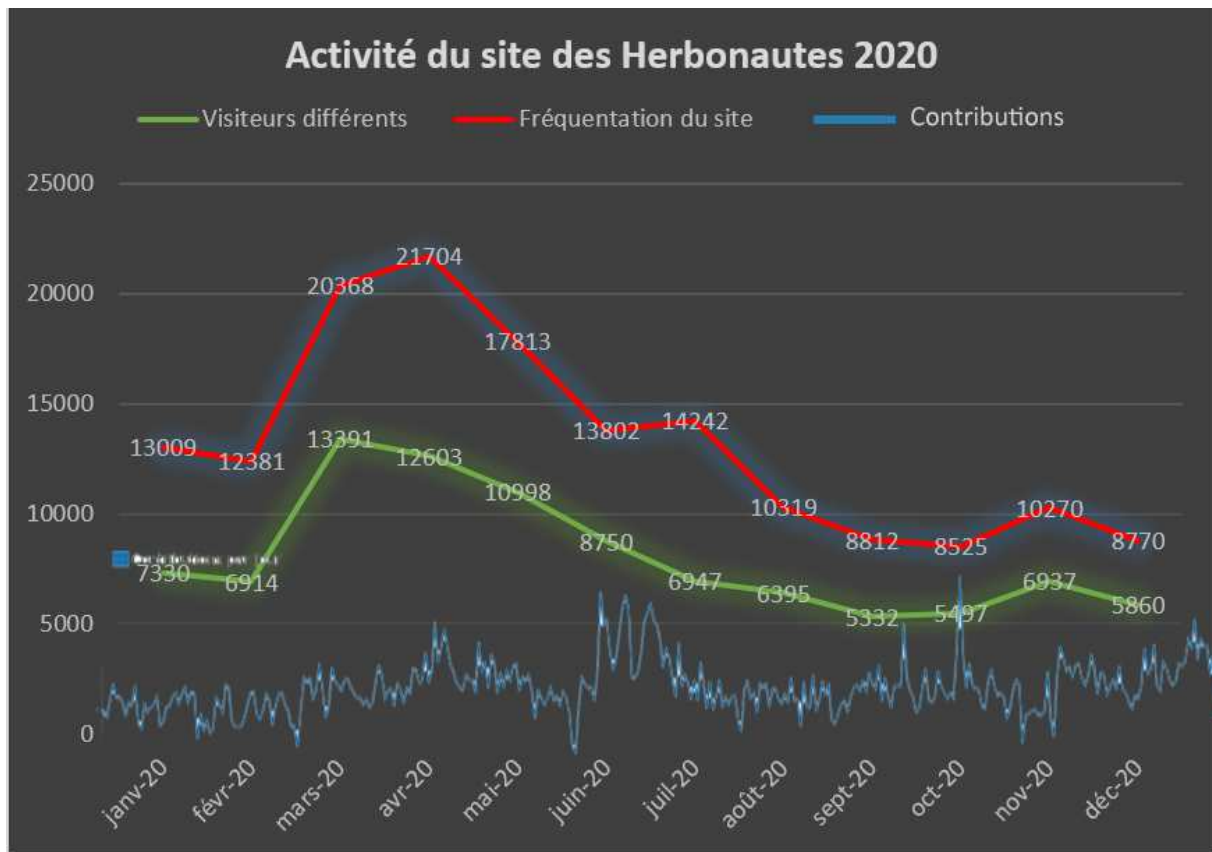
7571
spécimens complets

Avancement de la mission

Objectif : 7571 / 7571 specimens

Top contributeurs

Projet IHMAGES, les données de 7571 spécimens ont été informatisées par a communautés des Herbonautes



Activité du site des Herbonautes : une augmentation de la fréquentation a été enregistrée pendant le printemps 2020 (période de confinement dû à la crise sanitaire Covid 19).