



Annotate

Annotation rapide d'images et de vidéos
en sciences naturalistes et humaines

Séminaire IIIF

24 mars 2021

Laboratoire Dicen-IdF

www.dicen-idf.org

gilles.bertin@lecnam.net



Annotate pour les sciences naturalistes

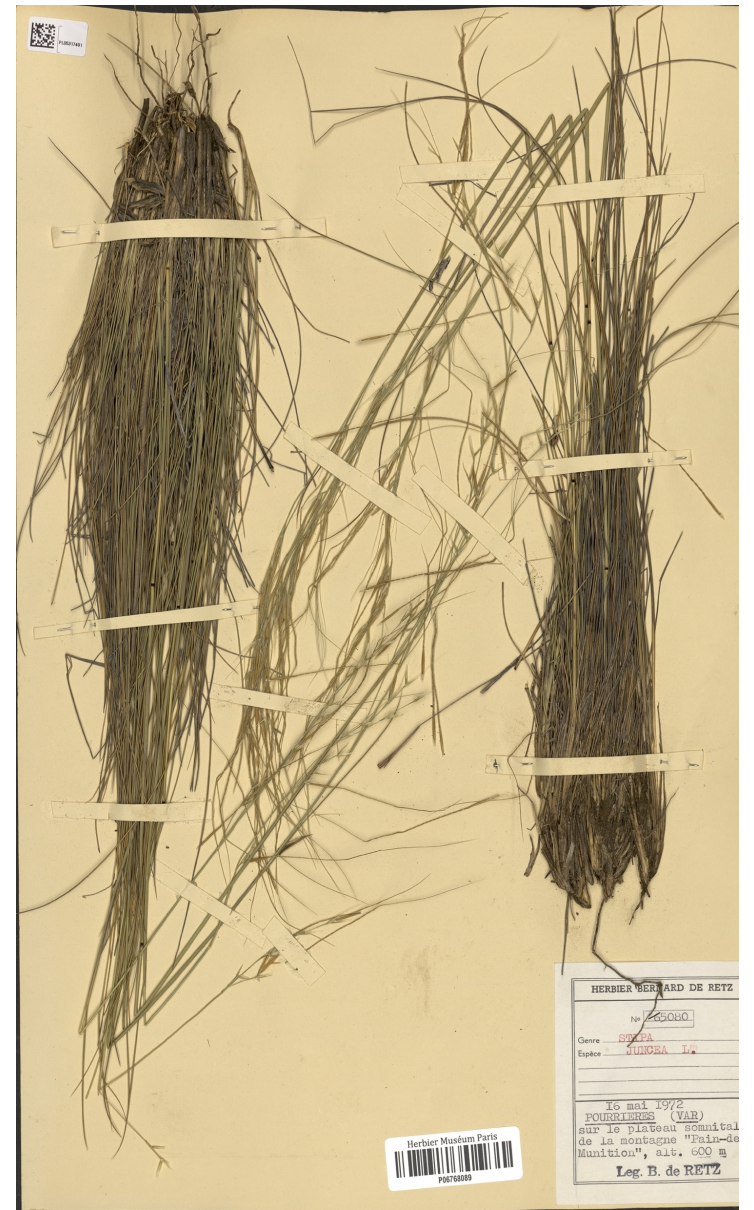
Projet ReCOLnat

L'infrastructure ReColnat, 10,9 millions d'images

- ANR ReColnat, devenue infrastructure ReColnat. Ses membres fondateurs :
 - MNHN (institution porteuse), Université Montpellier, Université Clermont Auvergne, Université de Bourgogne, IRD, INRA, Cnam, TelaBotanica, Agoralogie, GBIF, CNRS, ANR, Investissement d'avenir, Université Lyon 1
- Met à disposition les collections naturalistes numérisées des principales institutions françaises : 89 partenaires
- Au 17/2/2021 : 10 902 733 spécimens

Source chiffres : <https://www.recolnat.org/fr/les-chiffres>

The screenshot shows the ReColnat website homepage. At the top, there is a navigation bar with the ReColnat logo and links for ABOUT, EXPLORE, LESHERBONAUTES, ANNOTATE, and LAB. A red banner says "Bienvenue, sansaccent". Below this, a dark blue section titled "INFRASTRUCTURE RECOLNAT" displays statistics: 10 802 659 specimens, 4703 users, and 5 763 090 contributions. It includes a description of the virtual laboratory and links for "En savoir plus", "eReColNat en chiffres", and "Faire un don". A central image of a butterfly is labeled "EXPLORER LES COLLECTIONS". To the right, a green section titled "ACTUALITÉS" lists recent news items. Below these are sections for "NOS SERVICES, NOS OUTILS" (with links for "Mesurer et annoter vos spécimens", "Consulter nos spécimens", and "Plateforme de science participative Les Herbonautes") and "ENTRER DANS LE LABORATOIRE VIRTUEL" (with a link for "Commencez votre carrière dans le Collaboratoire").



Annotate, annoteuse collaborative rapide

- Annotations situées spatialement dans les images ([W3C annotation data model](#))
- Ergonomie propre aux métiers de la recherche naturaliste conçue pour l'annotation d'herbier et de collections en zoologie ou paléontologie, constitués d'images standardisées
- Contexte de tâches très homogènes et répétitives sur des dizaines ou centaines d'images, portant sur des projets de recherche variés : analyser des corpus de plusieurs centaines de planches et de constituer de tableaux de résultats d'observations utilisables dans des outils tiers en quelques heures

Annotate : plusieurs gammes d'outils

- Une palette de mesures physiques (longueur, angle) en sciences naturalistes.
- Une palette zones d'intérêts dans les images, dénombrement, transcription, ...
- Un tagueur d'images et d'annotations et un éditeur de métadonnées
- Une bibliothèque avec outils de sélection et d'agencement
- Un éditeur de modèles d'annotations permettant de gérer des caractères
- Un couplage avec Xper
- Un éditeur de collections virtuelles pour constituer des séquences d'images à exposer, écrire sur ces images, éditer leurs cartels, les exporter.
- Une API et un serveur Reolnat IIIF de présentation



Annotate en humanités numériques et archives

Projet OPAHH iiif

Open Pictures Annotator for Heritage and Humanities with iiif

Le projet OPAHH-iiif

«Open Pictures Annotator for Heritage and Humanities with iiif »

- Projet porté par le laboratoire Dicen-IdF au sein du Labex *Les passés dans le présent*, en réponse à un appel à projets 2019, avec
 - les Archives Nationales,
 - la BnF et son Département de l'audiovisuel,
 - et l'unité Esthétiques, sciences et technologies du cinéma et de l'audiovisuel de l'Université Paris 8.

Durée 3 ans, de 2020 à 2023

Le projet OPAHH-iiif

«Open Pictures Annotator for Heritage and Humanities with iiif »

Objectifs :

- offrir aux chercheurs en humanités numériques et professionnels du patrimoine (Musées, bibliothèques, archives) un outil ergonomique, contributif et ouvert pour faciliter, accélérer, approfondir leurs phases d'analyse d'images et de vidéos et rendre réutilisables leurs résultats par la communauté scientifique via le format d'échange IIIF ;
- proposer un outil standard, répondant aux besoins les plus fréquents de l'annotation patrimoniale et scientifique ;
- offrir une interface simple et intuitive à des non experts de l'annotation comme à des amateurs dans le cadre de projets participatifs ;
- explorer une application de la norme IIIF A/V à l'image animée.

Les attendus du projet

Exploration rapide de grands corpus d'images hétérogènes

Offrir à chaque chercheur en SHS la possibilité d'explorer rapidement des corpus d'images hétérogènes divers, y compris de photos numériques, en les annotant et en les taguant, puis en utilisant ces annotations dans des algorithmes ou outils tiers.

Éditorialisation

Permettre de sélectionner des images, au sein d'ensemble de corpus, de les légender, annoter, contextualiser pour en constituer des collections virtuelles exploitables dans des livres, vidéos, sites web.

Aller vers le IIIF

Expérimenter avec une chaîne IIIF existante, valider et mettre à disposition du Labex un service IIIF.

Deep learning

Créer rapidement à partir de corpus les données d'apprentissage nécessaires à la nourriture des algorithmes de deep learning applicables par exemple à des archives.

Ouverture

- Des outils de génération précédente conservent les annotations dans leurs bases de données internes. Il y a un enjeu fort désormais de pouvoir placer l'annotation dans une chaîne d'analyse ou de traitements.
- En entrée : lecture de sources d'images diverses, répertoires fortement hiérarchisés, images géolocalisées de téléphones, etc.
- En sortie : export des annotations en formats ouverts tableur, IIIF