

2024

www.info.persee.fr

Persée

RAPPORT D'ACTIVITÉ

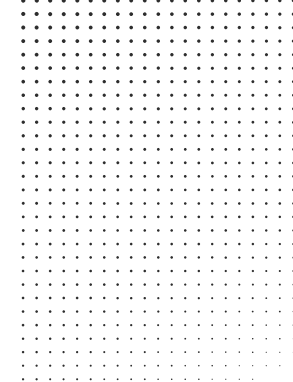




Illustration de couverture

Anemia hirta. Traité des fougères de l'Amérique (Plumier, 1705)

Cremers Georges, Boudrie Michel, Viane Ronald, Aymonin Gérard Guy.
« Le Père Charles Plumier (1646 – 1704) : son œuvre, son herbier de ptéridophytes américains. – 2. Typifications. »
In : *Le Journal de botanique*, n°103, 2022. pp. 2-117.



La mission historique de Persée, celle de numériser, enrichir et diffuser les publications scientifiques depuis leur origine, exige de faire face à la fois au temps long, aux problématiques de collecte et d'archivage, de garantie d'accès et de stabilité des formats, et à la fois aux évolutions technologiques sidérantes qui modifient les conditions de sa mise en œuvre et les pratiques de recherche qu'elle soutient.

Cette tension a largement traversé l'activité de Persée en 2024. **Comment faire évoluer** notre plateforme de production, nouer de nouveaux partenariats pour conduire ces évolutions, augmenter la qualité de nos données qui n'alimentent plus seulement un portail de diffusion mais sont de plus en plus mobilisées pour des usages de fouille, et poursuivre en parallèle le développement de l'offre éditoriale, en lien à la fois avec des éditeurs scientifiques et des bibliothèques de recherche toujours plus nombreux à vouloir rejoindre Persée ?

Ces objectifs ont engagé l'unité dans un **travail de redéfinition de ses axes de développement** dont la première phase commencée fin 2024 s'achève au printemps 2025. Ils ont guidé la **réorganisation de l'unité**, effective depuis janvier, pour mieux prendre en compte les questions de qualité, fédérer les compétences en traitement de données, et consolider nos projets en lien avec l'IA.

Une grande part de ces efforts accomplis en 2024 porte ses fruits en ce début 2025. Mais de premiers jalons ont déjà marqué l'année écoulée, tant au niveau des réalisations que de l'essor des coopérations qui permettent à Persée de jouer pleinement **son rôle d'opérateur national pour la valorisation du patrimoine scientifique** : l'amélioration de la qualité des PDF proposés aux lecteurs, le lancement de nos travaux dans le cadre de l'ÉquipEx Biblisisma+, la structuration du programme de numérisation enrichie de CollEx-Persée, que nous portons avec l'Université de Lille, l'entrée de Persée dans le Keepers Registry qui acte le rôle de l'UAR à un niveau international dans la conservation des publications scientifiques en sont quelques-uns.

Conjuguer stabilité et transformation n'est pas seulement un défi technique mais aussi humain. C'est grâce à l'énergie et l'ouverture au changement dont fait preuve l'équipe qu'aborder ces défis est aujourd'hui possible.

Gabrielle RICHARD
Directrice de Persée



Sommaire

	Produire et Ouvrir	03
	Explorer et Partager	11
	Données budgétaires	15
	Ressources humaines	21



Produire
& *ouvrir*

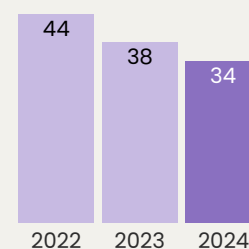
1 Sélection & intégration

En 2024, trente-quatre nouvelles collections ont été acceptées sur le portail Persée, représentant une volumétrie prévisionnelle de 315 000 pages. Les fonds éditeurs tiennent une place importante dans cette sélection, avec une première vague issue des Éditions Universitaires de Lorraine (11 collections) et des Presses Universitaires Blaise-Pascal (1 collection sélectionnée, 2 en préparation).

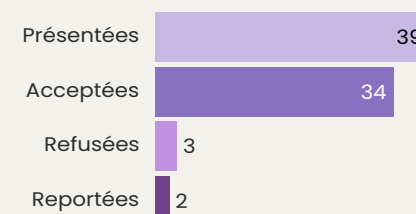
Côté Programme de Numérisation Concertée (PNC), 8 collections, appartenant aux Plans de Conservation Partagée des Périodiques (PCPP) de l'Enseignement supérieur, ont pu rejoindre Persée.

Candidatures au portail Persée sur 2 comités de sélection annuels

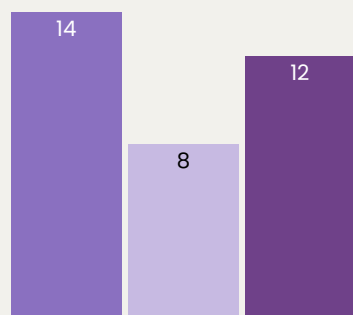
Évolution des candidatures



Détail des candidatures 2024



Origines des collections acceptées en 2024

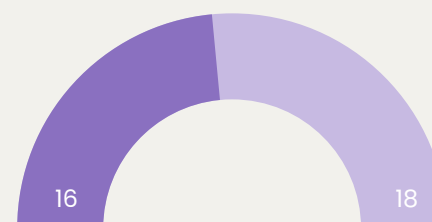


- Revue proposée par des éditeurs
- Revue et collections issues de PCP
- Collections issues de fonds éditeurs

Répartition livres / périodiques dans les collections acceptées en 2024

► En 2024, de nombreuses collections de livres issues de fonds éditeurs ont été sélectionnées.

- Collections de périodiques
- Collections de livres



2 Production & diffusion

Près de 470 000 pages ont été numérisées et plus de 525 000 pages documentées cette année, confirmant la capacité de Persée à maintenir une production soutenue malgré un contexte contraint (complexité des collections PNC et Perséides, chantiers techniques et qualité, poste vacant).

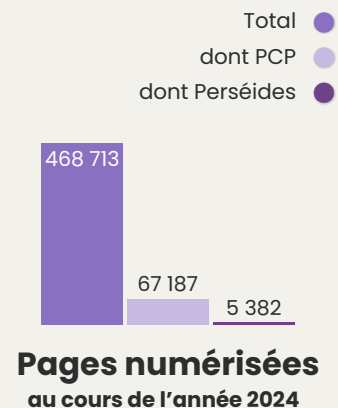
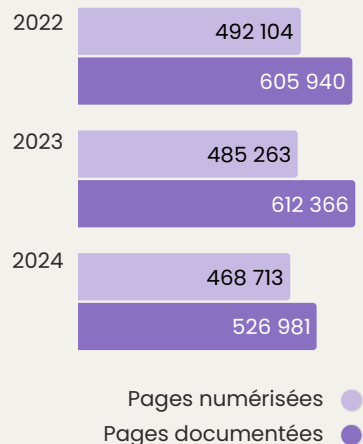
◆ Numérisation

La numérisation de la première collection de la (future) perséide *Magasin*, consacrée aux catalogues commerciaux et industriels à partir des années 1830, a nécessité une adaptation significative des processus de modélisation et de traitement pour répondre aux spécificités du matériel (organisation des contenus, dépliants recto-verso, finesse du papier, etc.).

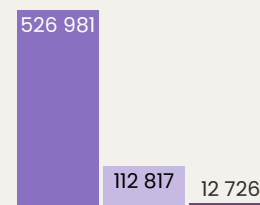
◆ Documentation

Côté documentation, la multiplication des partenariats externes a permis de compléter la production annuelle à hauteur de 38 743 pages, grâce à la contribution active du SCD de Lille, de la Bibliothèque interuniversitaire de la Sorbonne (BIS), de l'INHA et des Presses universitaires Blaise-Pascal.

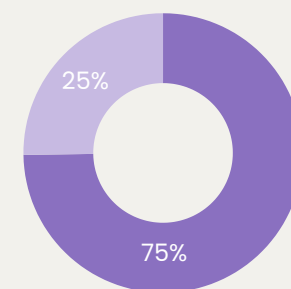
Évolution de la production



Pages documentées au cours de l'année 2024



Part du numérique natif dans les pages documentées en 2024



Si les collections imprimées restent majoritaires dans les corpus traités par Persée, les fichiers numériques natifs (PDF imprimeurs) ont compté pour 25% des pages documentées cette année.

Collections mises en ligne en 2024



21

nouvelles collections



103

collections complétées

◆ Diffusion

271 000 pages ont été mise en ligne en 2024 sur le portail Persée et les Perséides. Un chiffre en baisse par rapport à l'an dernier, du fait de l'important chantier d'amélioration de la qualité des PDF mené entre juillet et décembre 2024 (voir page 8), qui a reporté les nouvelles diffusion au début de 2025.

◆ Apport de CollEx-Persée

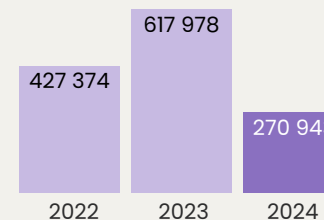
En 2024, le PNC a permis l'intégration de 10 nouvelles collections, qui viennent s'ajouter aux 17 collections issues des PCPP déjà diffusées. Cette avancée résulte d'une collaboration étroite avec les **bibliothèques partenaires**, telles que la BIS, Cujas et Henri Piéron (Université Paris Cité), qui ont contribué à réunir les collections et à contrôler la qualité des données diffusées. Ces mises en ligne marquent la finalisation de projets importants débutés dans le cadre du PNC de **CollEx-Persée 1**.

◆ Politique d'accès

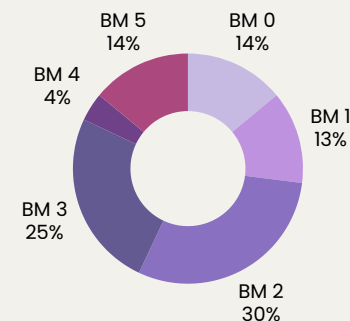
11 collections sont passées en barrière mobile cette année, portant à 100 le nombre de collections en diffusion continue, soit un peu moins d'un quart des collections présentes sur Persée.

8 de nos partenaires éditeurs ont par ailleurs opté pour une **réduction ou une suppression de la période d'embargo**, accélérant ainsi l'accès libre à leurs publications les plus récentes.

Évolution des mises en ligne en nombre de pages



Répartition des barrières mobiles au 31 décembre 2024



3 Données & métadonnées

◆ Nouveauté 2024 : la gestion des autorités collectives

Cette année Persée a implémenté **la gestion des autorités pour les organismes et collectivités** dans son système d'information.

Elle permettra de bénéficier du travail de **référencement des éditeurs accompli dans IdRef** par les partenaires du projet Mir@bel 2022, et de **simplifier la gestion des droits attachés aux images** selon les préconisations de l'« Étude sur l'utilisation d'œuvres relevant des arts visuels dans les publications scientifiques », commandée par le MESR.

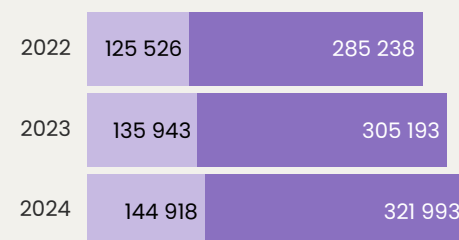
◆ Alignements sur des référentiels

Le travail documentaire mené en 2024 a permis une **nette augmentation des alignements d'espèces avec le référentiel GBIF**.

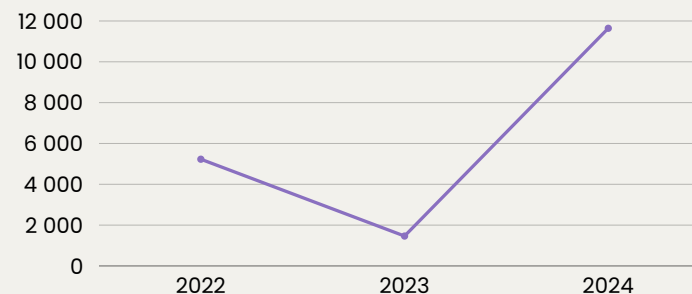
Cette hausse est principalement liée à la **nature des collections traitées** cette année, particulièrement riches en contenus taxonomiques, telles que les *Annales de la Société Géologique du Nord*, les *Mémoires de la Société d'Histoire Naturelle de l'Afrique du Nord*, les *Bulletins de l'Académie Vétérinaire de France*, ou les séries en sciences de la vie et de la terre des *Comptes Rendus de l'Académie des Sciences*.

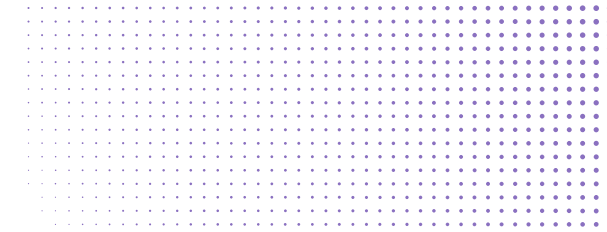
Autorités personnes gérées par Persée

● Total
● dont alignées sur IdRef



Nouveaux alignements d'espèces sur le référentiel GBIF





◆ Taux de reconnaissance OCR

La qualité de l'OCR (reconnaissance optique de caractères) constitue un enjeu majeur pour l'exploitation des corpus numérisés. Dans ce contexte, les résultats obtenus pour la production de 2024 sont particulièrement encourageants : **le taux de reconnaissance des caractères se situe entre 96,20 % et 96,45 %, et celui des mots entre 96,98 % et 97,52 %** (estimations par intervalle de confiance à 95 %).

Ces résultats se placent dans la **fourchette haute** de la catégorie "OCR moyen à bon", visée pour une meilleure ré-exploitation du texte.

Ils remettent en évidence la **puissance du moteur d'OCR actuel** et l'attention à porter aux éléments impactant la qualité du texte comme l'état initial du document, le multilinguisme, la présence d'éléments non textuels, les mises en page complexes.

◆ Persée intègre le Keepers Registry

Tout au long de l'année 2024, **Persée a préparé son entrée dans le Keepers Registry**. Ce service, géré par le Centre International ISSN, permet d'informer la communauté des bibliothèques et des éditeurs sur la conservation des publications en séries pourvues d'un ISSN (International Standard Serial Number).

Rejoindre les 19 institutions qui contribuent au Keepers Registry est **l'occasion de réaffirmer l'importance de l'archivage pérenne** dans les missions de l'unité et de donner une nouvelle visibilité à la part qu'elle prend dans cette initiative internationale, en terme de référencement, d'ouverture et de préservation des collections.

◆ Chantier "Qualité PDF"

Plusieurs facteurs ont conduit à la mise en place de ce chantier, notamment une dégradation de la qualité de certains PDF et le changement de licence d'une brique logicielle clé, auparavant Open Source.

L'opération a mobilisé l'équipe de manière transversale et extensive : **audit et diagnostic, développements informatiques, redéfinition des niveaux de qualité attendus et réalisés** (lisibilité, fidélité à l'original, taille).

Les **mises en ligne, suspendues sur plusieurs mois** au cours du second semestre 2024, ont pu reprendre en décembre sur la base d'une **nouvelle implémentation du processus de génération des fichiers PDF**, garantissant la conformité des fichiers produits.

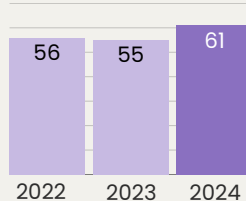
4 Usages & visibilité

Avec **61 millions de visites en 2024**, le portail Persée enregistre une fréquentation en **hausse de plus de 10 %** par rapport aux deux années précédentes. Au total, **88 millions de documents ont été consultés** au cours de l'année, soit 20 millions de plus qu'en 2023.

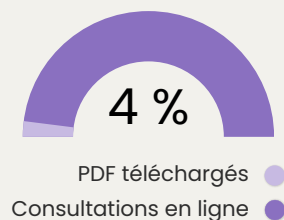
Les Perséides ont quant à elles généré **1,3 million** de consultations, participant ainsi activement à l'enrichissement et à la diversification de l'offre documentaire de Persée.

Le portail Persée

Évolution des consultations en millions de visites

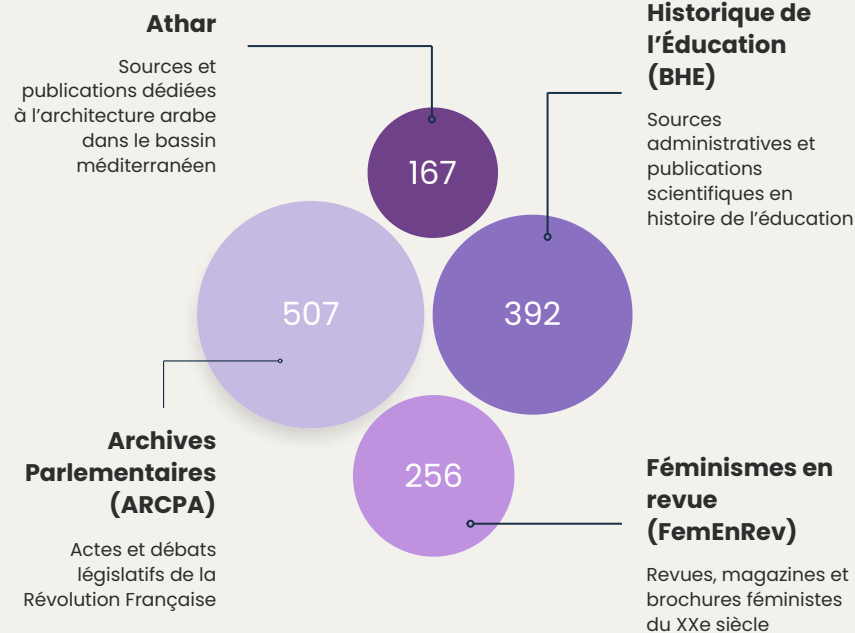


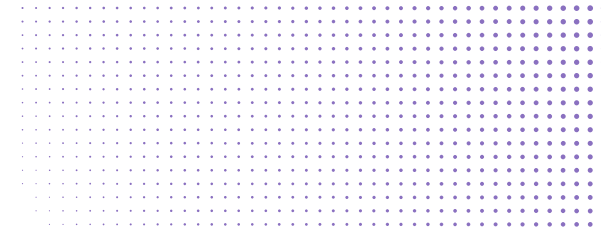
Part des téléchargements



Les Perséides

Consultations annuelles en milliers de visites



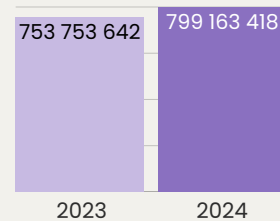


Près de 800 millions de pages ont été vues en 2024, soit une **augmentation de plus de 45 millions de pages** par rapport à 2023. Parmi les 10 pays les plus utilisateurs du portail, 4 se trouvent en Europe (France, Belgique, Italie, Suisse), 2 en Amérique du Nord (Canada, États-Unis) et 4 en Afrique (Maroc, Côte d'Ivoire, Algérie, Congo).

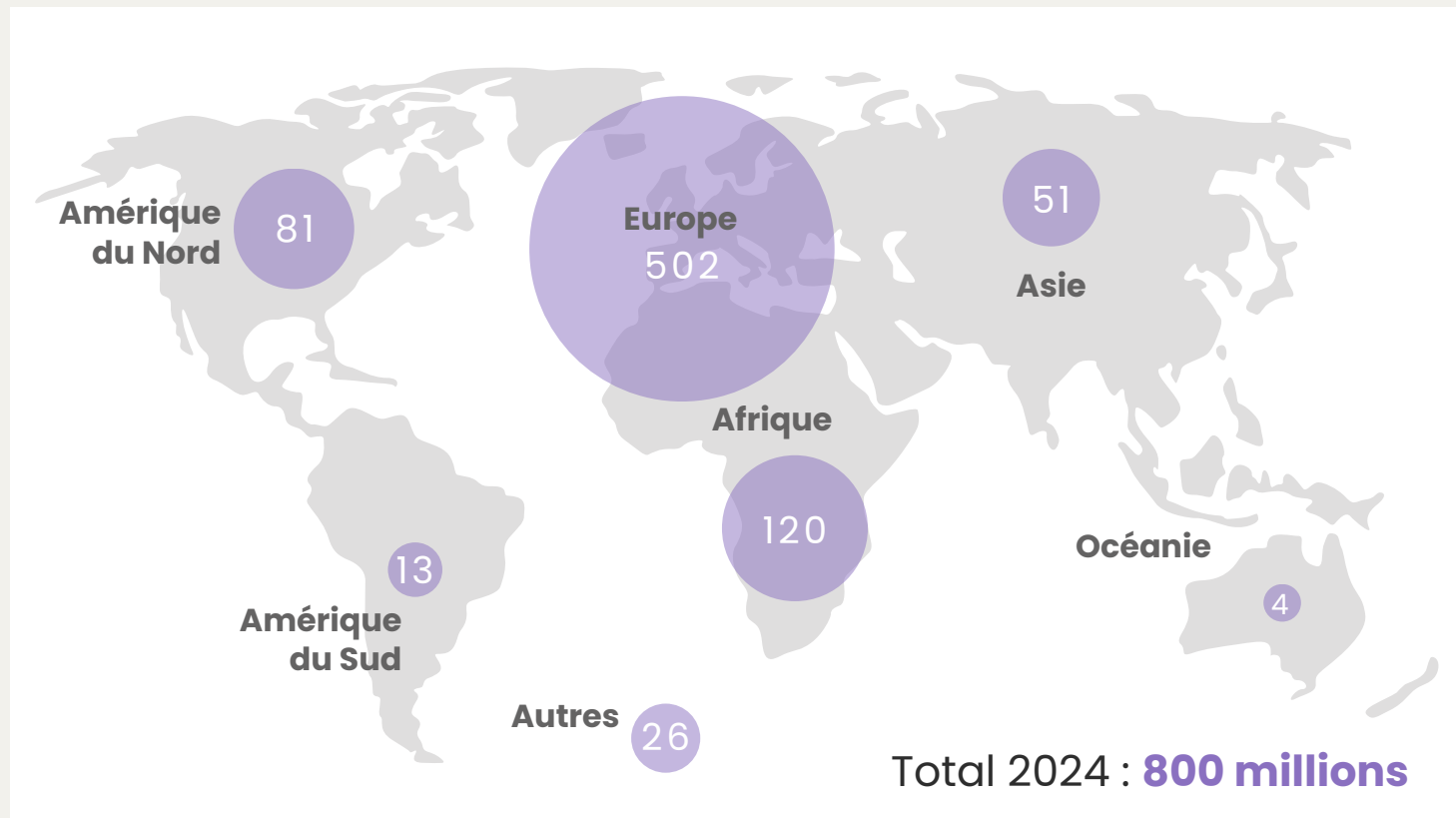
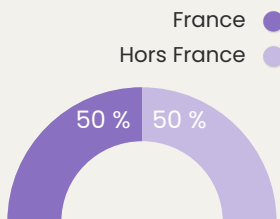
Le portail Persée

Consultations à l'international en millions de pages vues >

Évolution



Part des consultations hors France





Explorer & *partager*

1 Projets exploratoires

◆ Biblissima+

Dans le cadre de sa participation à l'**ÉquipEx+ Biblissima+**, Persée a entrepris de mettre en relation les sources primaires anciennes du portail Biblissima avec la littérature secondaire du portail Persée.

L'**application de techniques d'intelligence artificielle** rend possible la détection en masse des cotes de manuscrits dans les publications scientifiques et leur mise en correspondance avec le référentiel des cotes de Biblissima.



Le projet se poursuivra jusqu'à fin 2025 et a pour ambition d'enrichir les données des deux infrastructures et de produire des rebonds entre plateformes par le biais des identifiants pérennes des ressources, ici les articles présents sur Persée et les manuscrits référencés sur Biblissima. Le projet doit également permettre de mettre en place des mécanismes de traitement, de flux et de gestion de données durables.

Pour ce projet, **Persée a recruté un Chargé de traitement de données documentaires**, dont la mission est de contribuer à l'identification, à la structuration et à la mise en correspondance des données entre les deux plateformes.

◆ RAPIDO

En 2024, le projet Rapido a connu des avancées significatives. Un travail de fond a été mené sur le corpus de référence, avec la constitution d'un échantillon d'articles annoté et retranscrit manuellement, servant de base de comparaison pour **l'évaluation du modèle de détection automatique des toponymes**.

Le modèle de reconnaissance développé initialement par l'INIST a été considérablement amélioré, atteignant un **taux de performance de 97 % pour la détection automatique** des lieux et de **76 % pour le liage** avec les autorités IdRef.

Des **interfaces de validation** ont également été mises en place pour permettre l'implication directe des chercheurs dans le processus. Le projet se poursuivra en 2025, afin de consolider ces premiers résultats et d'explorer de nouvelles pistes d'exploitation scientifique.

2 Formations et collaborations

◆ Journées Omeka & NumaHOP

Persée a accueilli à l'ENS de Lyon les **journées Omeka & NumaHOP**, qui se sont tenues en novembre.

Cet événement a rassemblé des professionnels de l'information, des chercheurs et des institutions autour des **deux plateformes open source** et des problématiques de production et de diffusion de données issues de la numérisation.

Cette première édition de journées communes aux deux associations d'utilisateurs s'inscrit dans la démarche de **promotion et de développement de solutions libres pour la gestion et la valorisation du patrimoine scientifique numérisé** que Persée porte au sein du GIS.



Affiche de l'évènement >
réalisée par Persée

◆ Partage d'expertise

Les initiatives de partage d'expertise se sont poursuivies sur la lancée de l'année 2023, à travers **des formations** (master Informatique de l'Université Lumière Lyon 2, master Humanités Numériques de Lyon) **et des participations à des journées thématiques** à destination des communautés de recherche et d'ingénierie, telles que les journées Biblissima+, le colloque "Les Presses universitaires et l'édition numérique : bilan et perspectives" (projet FNSO PEPSO-BFC) ou le séminaire GANESHs (Groupe d'Action sur le Numérique en Sciences et Humanités, Université Jean Monnet Saint-Étienne).

◆ Numérisation enrichie

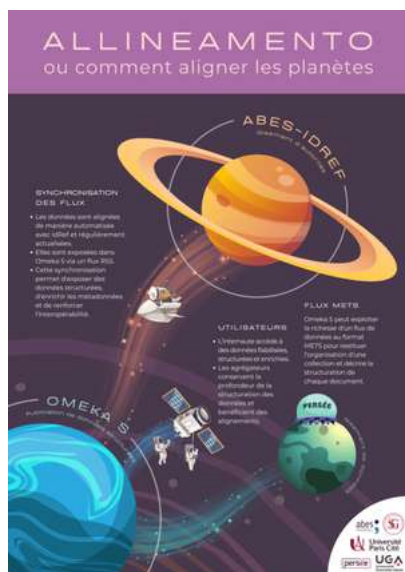
Le **groupement d'intérêt scientifique CollEx-Persée**, qui lie Persée à un réseau de bibliothèques de recherche et d'opérateurs de l'information scientifique, a été renouvelé en octobre 2024. Au sein du GIS, Persée copilote avec l'Université de Lille le programme dédié à la **numérisation enrichie**.

Trois axes structurent le programme : l'animation d'un réseau professionnel dédié à la production et l'exploitation des corpus documentaires numériques, l'ouverture et l'interopérabilité des outils et des données, une politique concertée de numérisation de la littérature scientifique.

◆ Allineamento : un poster primé

Le poster *Allineamento*, fruit d'un travail commun entre l'Université Paris Cité, la Bibliothèque Sainte-Geneviève, l'Université Grenoble-Alpes (UGA) et Persée, a été récompensé par une **3e place lors des Journées de l'Abes 2024**.

Ce projet fait partie des sept initiatives sélectionnées dans le cadre du programme de coopération de l'Abes 2023. Il vise d'une part à intégrer plus étroitement le **référentiel IdRef** au sein de la plateforme libre **Omeka**, afin de renforcer la qualité, l'enrichissement et l'interopérabilité des métadonnées des collections numériques. Et d'autre part à **exploiter dans Omeka S un flux de données au format METS**, tel que celui que produit Persée, afin de restituer l'organisation d'une collection.



◀ Poster présenté aux JABES 2024

◆ DATAlystE : vers une labellisation en 2025

Lancé dans le cadre de l'écosystème **Recherche Data Gouv**, DATAlystE est un **atelier de la donnée** dont la mission est d'accompagner les chercheurs, doctorants et personnels d'appui dans la **gestion de leurs données de recherche**, selon les **principes de la science ouverte et des données FAIR**.

Porté par un **collectif de 12 partenaires** académiques et institutionnels du site Lyon Saint-Étienne, dont Persée, DATAlystE a franchi une étape décisive en 2024 à travers la définition de ses instances et de son offre de service en vue d'une **labellisation en 2025**.

Son site web, articulé autour d'un guichet unique de ressources locales et nationales, et proposant un parcours de la donnée et un agenda partagé, constitue désormais un point d'entrée central pour accéder à ces services.



Site web de DATAlystE : <https://datalyste.universite-lyon.fr/>

L'atelier s'est également engagé activement au sein du **réseau national des Ateliers de la Donnée (ADLD)**, notamment par sa participation à plusieurs groupes de travail thématiques.



Données *budgétaires*

1 Recettes globalisées

TUTELLES	2022	2023	2024	2025
MESR – dotation initiale	72 000 €	72 000 €	72 000 €	72 000 €
MESR – dotations exceptionnelles	50 000 €	25 000 €	100 000 €	
ENS – dotation initiale	41 100 €	37 000 €	46 661 €	40 000 €
ENS – dotations exceptionnelles		-11 000 €	11 000 €	
CNRS – dotation initiale	35 000 €	35 000 €	33 957 €	35 000 €
CNRS – dotations exceptionnelles	100 000 €			
TOTAL	298 100 €	158 000 €	263 618 €	147 000 €

Fléchage des dotations exceptionnelles :

2022

MESR et CNRS : Renouvellement de l'infrastructure (site Monod)

2023

MESR : Prestation de conseil (à ouvrir en 2024)

ENS : Prélèvement

2024

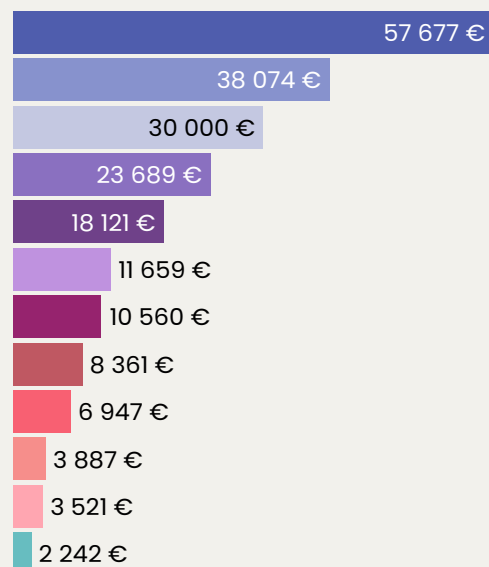
MESR : Investissement (scanners)

Libération des images (à ouvrir en 2025)

ENS : Restitution du montant prélevé en 2024

2 Dépenses globalisées

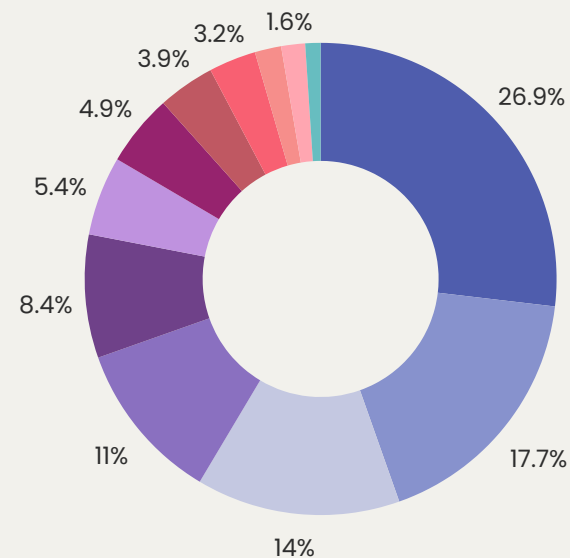
Dépenses 2024
sur recettes globalisées



- Numérisation
- RH sur ressources propres
- Développement stratégique
- Matériel informatique
- Archivage au CINES
- Missions
- Formation et alternance
- Communication et réception
- Frais généraux
- Gestion matérielle des collections
- DOI CrossRef
- Adhésions professionnelles

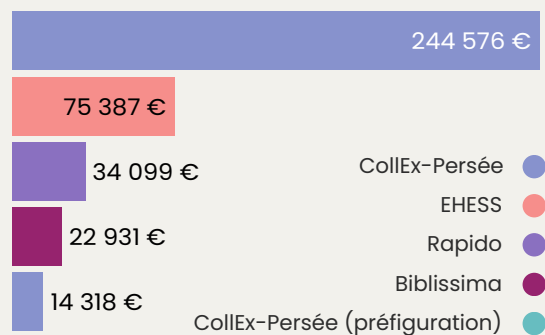
Total
AE consommées
214 738 €

Répartition



3 Dépenses fléchées

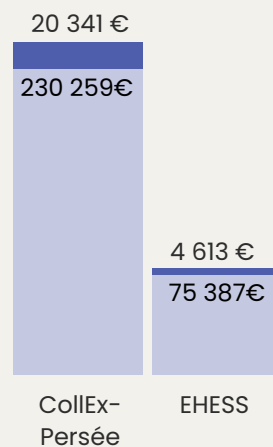
Dépenses 2024 sur recettes fléchées



Total
391 311 €

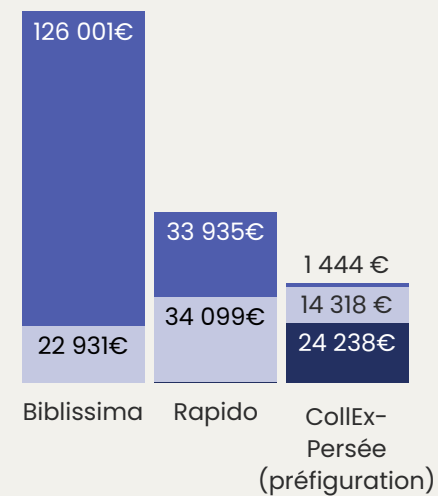
Contrats avec subventions annuelles

AE consommées en 2024
Montant restant



Contrats pluri-annuels

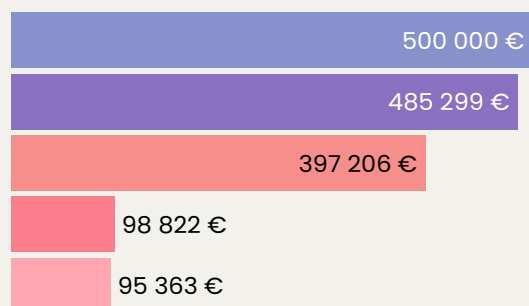
AE consommées en 2024
Consommé au 31/12/2023
Montant restant



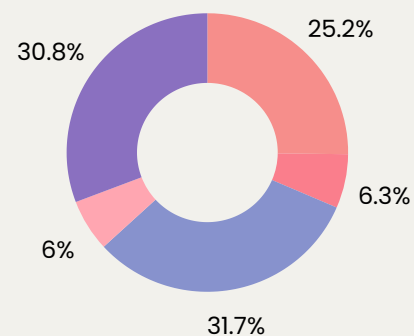
4 Masse salariale

- Postes affectés à l'ENS pour Persée
- Postes créés à l'ENS pour Persée
- Subvention MESR pour postes contractuels
- Reste à charge ENS pour postes contractuels
- CNRS - Titulaires + contractuels

Masse salariale 2024



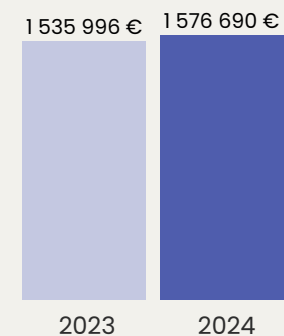
Répartition



Total

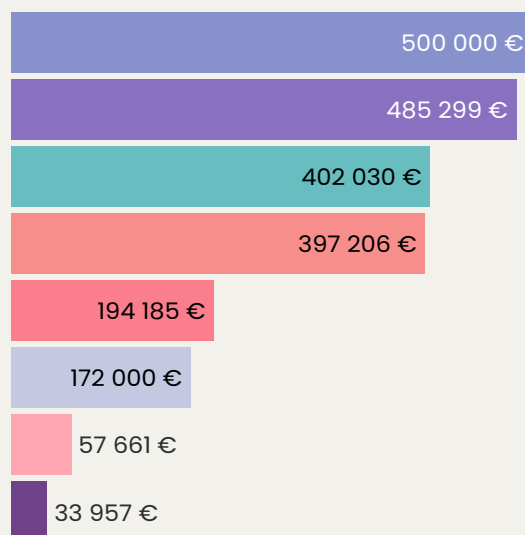
1 576 690 €

Évolution



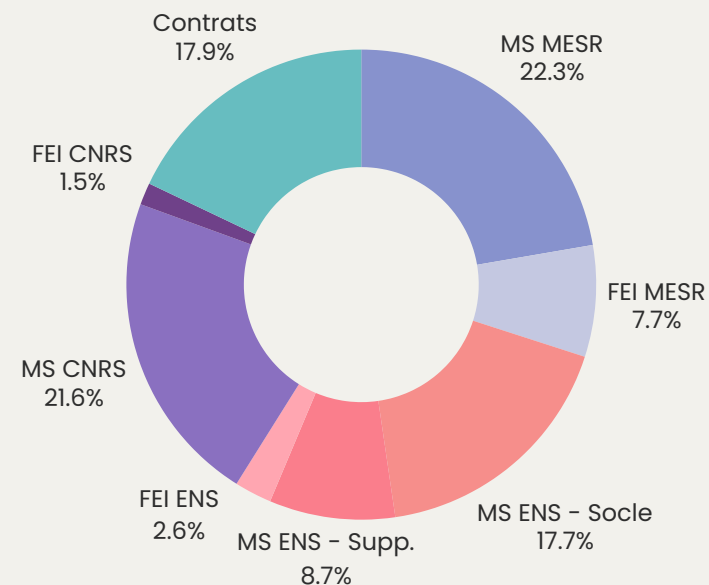
5 Total des recettes

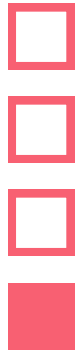
Origines des ressources masse salariale comprise



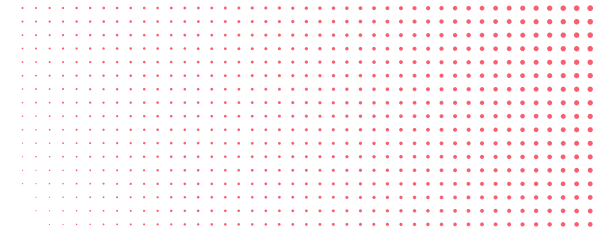
Total
2 242 338 €

Répartition



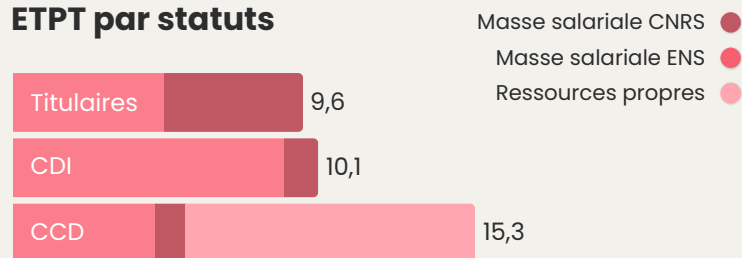


Ressources ***Humaines***

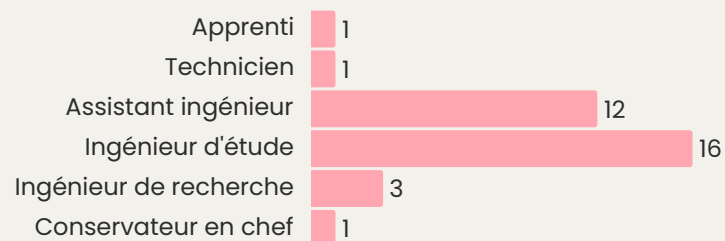


1 Répartition & mouvements

ETPT par statuts



Répartition par grades



Mouvements de l'équipe au cours de l'année 2024

5 départs
de l'unité



effectif 34
au 31 décembre



2 arrivées
dans l'année



● Hommes ● Femmes



2 Équipe Persée

au 31 décembre 2024

Direction

Gabrielle Richard
Directrice de l'unité

Missions rattachées à la Direction

Charlotte Fleytout
Gestion financière et
administrative

Sarah Gauthé
Cheffe de projet CollEx

Patricia Millier
Communication

Sitthida Samath
Qualité &
interopérabilité

Production

Éric Astier
Responsable du pôle

Sarah El Dirani
Numérisation

Nicolas Mahy
Numérisation et
post-traitements

Ludivine Poux
Numérisation

Documentation

Thomas Mansier
Responsable du pôle

Cécile Almonté
Documentation

Marie-Hélène Barrière
Publications numériques

Thibault Girard

Publications numériques
Support utilisateur

Magdalena Kruszynska
Documentation

Jenna Lagha
Documentation

Bérénice Magnant
Documentation

Flore Mayvial
Documentation &
Autorités

Blaise Moulin
Documentation

Loris Ohanian
Documentation

Savanah Samath
Documentation

Marie Vial-Bonacci
Documentation

Partenariats

Eugénie Michel
Partenariats
Bibliothèques (PNC/PCP)

Étienne Ferrandi
Traitement de données

Agnieszka Halczuk
Partenariats Recherche

Patricia Millier
Partenariats éditeurs

Samantha Saïdi
Partenariats Recherche

Informatique et développements

Viviane Boulétreau
Responsable du pôle

Vincent Beauseigneur
Développement

Émile Bouvier
Développement

Sylvain Boschetto
Développement

Nicolas Houillon
Développement

**Églantine Kremer-
Cochet**
Système et Réseaux
(alternance)

Clément Matray
Systèmes et Réseaux

Carole Vinot
SID

Ludovic Walle
Développement



info.persee.fr / www.persee.fr



partenariats@persee.fr



Persée - ENS de Lyon, 15 parvis René Descartes
BP 7000 - 69342 Lyon Cedex 07



