



**INIST**

**INSTITUT DE L'INFORMATION  
SCIENTIFIQUE ET TECHNIQUE**

**RAPPORT  
D'ACTIVITÉ  
2022**

Accompagnement et services  
numériques pour la recherche





À TRAVERS SON OFFRE  
DE SERVICES,  
L'INIST AFFIRME SON  
ENGAGEMENT DANS  
LE MOUVEMENT  
INTERNATIONAL DE LA  
SCIENCE OUVERTE.

+++  
+++  
+++

+++  
+++  
+++  
+++

+++  
+++  
+++

+++  
+++  
+++  
+++

+++  
+++  
+++  
+++

# SOMMAIRE

ÉDITO.....	p. 5
<b>Permettre aux communautés de la recherche d'accéder facilement à la production scientifique internationale.</b>	
PORTAILS ET PLATEFORMES.....	p. 8
NÉGOCIATIONS ET ACQUISITIONS.....	p. 10
HAL-CNRS.....	p. 12
ÉDITION NUMÉRIQUE.....	p. 13
<b>Faciliter l'accès aux données, leur interopérabilité ainsi que leur réutilisation.</b>	
OPIDOR.....	p. 16
DORANUM.....	p. 18
EXPOSITION DES DONNÉES.....	p. 20
<b>Répondre aux besoins de pilotage scientifique et d'extraction de connaissances en proposant une offre de services d'analyse et de fouille de textes.</b>	
DES RESSOURCES POUR LE TEXT & DATA MINING (TDM)...	p. 24
▪ Istex	
▪ Loterre	
INGÉNIERIE TERMINOLOGIQUE.....	p. 26
TEXT & DATA MINING (TDM).....	p. 27
ANALYSE ET VISUALISATION.....	p. 28
TRADUCTION.....	p. 29
<b>L'Inist, c'est aussi...</b>	
COMMUNICATION.....	p. 31
VEILLE.....	p. 31
WEB.....	p. 31
SECRÉTARIAT GÉNÉRAL.....	p. 32
SYSTÈMES D'INFORMATION.....	p. 32
LES RENCONTRES.....	p. 33



# ÉDITO

## ANNÉE DE CONFIRMATION POUR DE NOMBREUX SERVICES

Symbole de l'engagement de l'ensemble des membres de l'Inist pour **servir au mieux les communautés scientifiques**, l'équipe en charge du portail BibCnrs a été mise à l'honneur avec la remise d'un **Cristal collectif 2022** pour le déploiement de l'outil « **Click & Read** ». Ce dernier permet en un clic d'accéder au texte intégral d'un article quelle que soit la page web consultée et aussi quel que soit le statut de l'article : en accès libre depuis des sources diverses ou comme ressource électronique acquise par l'établissement. Si l'outil a d'abord été configuré pour accéder aux ressources négociées par le CNRS, **les collaborations se sont multipliées** toute l'année pour permettre l'accès aux ressources documentaires de 22 établissements partenaires.

Cette collaboration avec l'ensemble des établissements d'enseignement supérieur et de recherche est également centrale dans le développement des outils de mesure des usages de leurs plateformes documentaires. Toujours en partenariat étroit avec Couperin.org, **le nouvel outil ezCOUNTER** permet d'analyser les rapports d'utilisation fournis par les éditeurs des ressources documentaires et complète les outils déjà disponibles sur la plateforme commune ezMESURE.

Avec l'ouverture en juillet 2022 de la plateforme nationale fédérée des données de la recherche « **Recherche Data Gouv** » qui s'appuie sur un écosystème d'accompagnement des communautés scientifiques dans la gestion et le partage de données de recherche, l'Inist a accentué son implication dans l'administration et la curation de l'espace générique. Pour accompagner les ateliers de la donnée dans leurs missions, l'Inist porte également les **deux centres de ressources OPIDoR et DoRANum** permettant ainsi de mutualiser les outils et ressources pédagogiques indispensables dans la mise en œuvre de leurs missions.

Depuis 10 ans, l'Inist est **agence d'attribution de DOI** pour les jeux de données en France au sein du consortium international DataCite. Afin de dynamiser cette activité et définir une stratégie commune au niveau national, le consortium DataCite France se structure, sa **première assemblée générale** s'est déroulée le 1<sup>er</sup> décembre 2022 grâce à l'implication de toute l'équipe.

Dans le cadre de son service d'exposition des données, l'Inist déploie des **bibliothèques numériques** dédiées aux données de recherche pour le compte de 18 laboratoires et 5 MSH.

La collaboration avec la MSH Lorraine dans la mise en place de la **plateforme Cenhtor** est exemplaire dans la mesure où elle permet de définir un modèle de collaboration avec la MSH qui assure l'accompagnement des projets de recherche pour l'exposition et le partage de leurs données.

Dans le domaine de l'analyse et de la fouille de texte, cette année était également l'occasion de fêter les **10 ans d'existence de la plateforme Istex**. Dans ce cadre, le site a fait peau neuve en intégrant l'ensemble des fonctionnalités permettant d'exposer, de constituer et de traiter des corpus documentaires. Les nouveaux services de fouille de textes sont quant à eux ouverts aux utilisateurs sous la forme de **web services** disponibles sur le site Objectif TDM. La convergence de ces nouveaux services, des terminologies et des outils de visualisation comme Lodex est étudiée dans le **projet d'infrastructure de recherche** qui vise à développer les ressources documentaires et rendre accessible leur usage dans un cadre de fouille de textes.

L'Inist a participé activement aux travaux préparatoires à l'évaluation du CNRS, nous pouvons noter par exemple l'**étude bibliométrique** permettant de mettre en avant la place du CNRS dans les recherches portant sur l'évolution du climat et l'impact des activités humaines telles qu'elles sont citées dans les derniers rapports du GIEC.

L'ensemble des travaux décrits dans ce rapport d'activité sont réalisés en partenariat avec les opérateurs comme l'ABES, Couperin.org, Persée ou le CCSD, les collègues des services communs de documentation, le réseau des URFIST et également dans le cadre de projets comme le **projet TyDI** avec l'INRAE, le **projet ANR MATOS** porté par le Laboratoire interdisciplinaire des sciences du numérique de l'Université Paris Saclay et enfin les **projets EOSC**. Ceci place l'Inist au cœur d'un réseau national d'acteurs très riche en complément d'un travail étroit avec la Direction des données ouvertes de recherche (DDOR) comme acteur de la **politique Science ouverte du CNRS**.

Bonne lecture.

**Claire FRANÇOIS**  
Directrice de l'Inist

Permettre aux  
communautés de la  
recherche d'accéder  
facilement à la production  
scientifique internationale.



## PORTAILS ET PLATEFORMES

- ▶ Plusieurs réservoirs d'information incontournables
- ▶ Une extension web pour un accès simplifié

## NÉGOCIATIONS ET ACQUISITIONS

- ▶ Des ressources au plus près des besoins

## HAL-CNRS

- ▶ La vitrine de la production scientifique du CNRS

## ÉDITION NUMÉRIQUE

- ▶ Un soutien à l'édition scientifique

# PORTAILS ET PLATEFORMES

Grâce à des plateformes et outils, les publications scientifiques deviennent accessibles en quelques clics.

## CLICK & READ

**L'EXTENSION CLICK & READ** facilite l'accès aux publications scientifiques en affichant un **bouton cliquable directement sur la page du navigateur Internet** lorsque le document PDF est consultable.

L'outil parcourt la page Internet visitée à la recherche d'identifiants documentaires (DOI).

Lancée en 2021, Click & Read évolue en 2022 pour devenir **plus collaborative**.

- ▶ **PLUS GRANDE MIXITÉ DE RESSOURCES CNRS ET ESR.**
- ▶ **POSSIBILITÉ D'ORDONNANCER L'ENSEMBLE DES SOURCES.**

L'ajout de nouvelles sources a d'abord concerné des établissements qui utilisent les outils de la société Ebsco, puis s'est progressivement élargi aux sources des établissements qui utilisent les outils de la société ExLibris.

**18 WEBINAIRES DE PRÉSENTATION REGROUPANT PLUS DE 550 PARTICIPANTS**

**PLUS DE 20 000 INSTALLATIONS**

**435 871 CONSULTATIONS D'ARTICLES VIA CLICK & READ**

**Fin 2022, Click & Read consulte les 26 fonds documentaires suivants :** Unpaywall, Istex, Panist, BibCnrs, Insermbiblio, Université de Lyon 1, Université d'Orléans, Université de Rennes 1, Université de Limoges, CEA, B2RM, Institut Pasteur, Université Paul-Valéry Montpellier 3, Université de Montpellier, Université de Nîmes, Université Littoral Côte d'Opale, Sciences Po, Université de Bordeaux, Sorbonne Université, Collège de France, Université de Pau et Pays de l'Adour, BU Sainte Geneviève, Université Savoie Mont Blanc, Université des Antilles, Université Paris-Saclay, Université Jean Monnet Saint-Etienne.

## CLICK & READ, LAURÉAT DU CRISTAL COLLECTIF DU CNRS 2022

Le 14 novembre 2022, l'équipe Click & Read s'est vue remettre le **crystal collectif du CNRS 2022** à l'occasion de la cérémonie des Talents CNRS de la délégation Centre-Est qui s'est déroulée à Nancy.

C'est une belle reconnaissance pour l'équipe. Grâce à leur investissement dans la création de l'extension Click & Read, l'accès aux documents se fait plus simplement, de quoi faciliter le travail des chercheurs mais aussi de l'ensemble de la communauté de l'enseignement supérieur et de la recherche.



De gauche à droite. 1<sup>er</sup> rang : Sandrine Chopot, Dolorès Dardaine, Barbara Negre y Rossello, Laurence Grand. 2<sup>e</sup> rang : David Castelain, Florence Join, Martial Luc, Sylvain Ponticelli, Mounir Habsaoui.  
© Philippe Fournerey / Inist-CNRS

## BIBCNRs

### S'ADAPTER AUX BESOINS DES USAGERS DANS UN CONTEXTE DE SCIENCE OUVERTE

L'année 2022 a essentiellement été consacrée à des études préparatoires en vue de faire évoluer BibCnrs dès 2023 vers de nouveaux services :

- faisabilité d'ajout d'un **accès vers des données de la recherche** (basé sur le service interne MÉTADoRe qui moissonne le réservoir Datacite) ;
- réflexion sur la possibilité de mise en avant du **modèle de publication en accès ouvert Diamant** soutenu par le CNRS.

**bib**  
cnrs

5 134 749\* consultations en 2022 dont :

**2 464 099\*** articles scientifiques sous licences éditeurs

**145 718\*** articles scientifiques issus de ressources libres

**1 101 424\*** de recherches via des bases de données (hors consultation du Web of Science, non analysable par ezPAARSE)

*\*données ezPAARSE*



### LES UTILISATEURS

**54 143** chercheurs ayants droit

- **21 606** chercheurs CNRS
- **32 537** enseignants-chercheurs

### LEUR APPARTENANCE



**1523** unités CNRS tous types confondus



**8 389** abonnés aux listes de diffusion



**1 360** abonnés au compte Twitter

## Lancement du nouveau site OpenBibArt

### UN SIÈCLE DE BIBLIOGRAPHIE SUR L'ART OCCIDENTAL DÉSORMAIS DISPONIBLE EN LIGNE

Mise en ligne en octobre, OpenBibArt est une base de données qui recense **plus de 1,2 million de notices bibliographiques** sur des sujets liés aux arts et à l'archéologie publiés entre 1910 et 2007.

Elle est le fruit d'une collaboration fructueuse entre trois institutions : l'**Institut national**

d'**histoire de l'art** (INHA) à Paris, le **Getty Research Institute** (GRI) à Los Angeles et l'**Inist**.

Plateforme de recherche ouverte développée par l'Inist, OpenBibArt unifie le contenu de plusieurs ensembles de données connexes :

- Le Répertoire d'art et d'archéologie (RAA) ;
- Le Répertoire international de la littérature d'art/International Repertory of the Literature of Art (RILA) ;
- La Bibliography of the History of Art (BHA).

## Alimentation de la base INIS

L'Inist héberge et alimente pour le CEA une des plus grandes bases de données bibliographiques sur les publications mondiales en sciences nucléaires : **INIS** (International Nuclear Information System).

La base INIS compte plus de 4 millions d'instances et s'enrichit chaque mois de plusieurs dizaines de milliers de documents.

# NÉGOCIATIONS ET ACQUISITIONS

Toutes les acquisitions de ressources thématiques pour le portail BibCnrs sont gérées à l'Inist. Les négociateurs travaillent chaque année avec une cinquantaine d'éditeurs et diffuseurs internationaux pour le renouvellement ou la mise en place de nouveaux abonnements.

## ACHATS

### RENOUVELLEMENT 2023-2026 DE L'ACCORD-CADRE CNRS

Les **prestations d'acquisition et de gestion de documentation physique et électronique** font l'objet d'un accord-cadre porté par le CNRS, en collaboration avec la Direction déléguée aux achats et à l'innovation. Le besoin a été mis à jour, le cahier des charges a été réécrit et les modalités de réponse à l'appel d'offres ont été révisées.

Les prestations prévues dans cet accord-cadre sont accessibles à tout établissement de l'ESR français, dans un **esprit de mutualisation**.

## DOCUMENTATION ÉLECTRONIQUE

### CAMPAGNE DE DÉCOUVERTE

Afin de servir la communauté scientifique au plus près de ses besoins, une **campagne de découverte** est organisée chaque année. Elle permet aux chercheurs de tester en « conditions réelles », des revues, des bases de données, des livres électroniques.

En 2022, près de **20 ressources** ont été proposées à l'essai.

### ACQUISITIONS

Résultant de la campagne de découverte 2021, l'Inist compte de nouvelles acquisitions :

- les **revues de l'Electrochemical Society (ECS)** pour INC, INP, IN2P3, INSIS, INS2I ;
- **10 revues Wiley supplémentaires** sur les thématiques environnement/anthropologie, dont *Methods in Ecology and Evolution*, *Functional Ecology* et *Evolutionary Anthropology* ;
- un **abonnement expérimental** d'un an aux revues Cambridge University Press ainsi qu'à la newsletter TheMetaNews.

## RECENSEMENT DES FRAIS DE PUBLICATION CNRS

**LES DÉPENSES POUR 2018 DU CNRS** en frais de publication (APC) payés par les laboratoires pour publier en accès ouvert ont été recensées permettant à la direction du CNRS d'avoir des données quantifiées.

**LES DONNÉES 2018 ONT ÉTÉ FOURNIES À COUPERIN.ORG** pour être chargées dans OpenAPC.

# UN NOUVEL OUTIL DÉDIÉ À LA MESURE DES USAGES

## EZCOUNTER

### UN MODULE DANS EZMESURE POUR LA COLLECTE DES DONNÉES COUNTER 5

Développé en partenariat avec Couperin.org, ezCOUNTER est un **module spécialisé dans le moissonnage automatique des rapports d'utilisation** fournis par les éditeurs de ressources électroniques.

Il permet de **collecter massivement et périodiquement** les statistiques d'usage depuis 250 points d'accès certifiés COUNTER 5 et de visualiser l'utilisation des ressources à différents niveaux : titres, bases de données, plateformes, items. Ce développement a pour vocation de faciliter la conduite des négociations et la réponse aux enquêtes nationales telles que eSGBU (enquête statistique générale auprès des bibliothèques universitaires) et sa section réservée aux usages.

#### ezCOUNTER trouve son public

Dès la première année, **92 établissements partenaires** ont déclaré les identifiants SUSHI des plateformes de documentation électronique auxquelles ils sont abonnés.



## EZUNPAYWALL

ezUNPAYWALL, l'**API miroir d'Unpaywall** opérée par l'Inist pour le CNRS, s'est dotée d'une interface web facilitant l'interrogation et permettant l'enrichissement de fichiers.



Christine Weil-Miko, responsable des négociations nationales à l'Inist, a été élue le 29 novembre au **Bureau professionnel du consortium Couperin.org** au titre de responsable du Département des négociations documentaires.

# HAL-CNRS

Vitrine de la production scientifique du CNRS, le portail HAL-CNRS voit sa feuille de route validée et une équipe nouvellement constituée œuvrer pour son application.



## PORTAIL HAL-CNRS : NOUVEAU LOOK

Il s'inscrit dans le déploiement de la nouvelle ergonomie des interfaces de HAL impulsé par le CCSD.

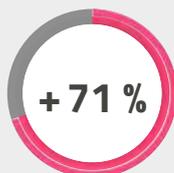
## UNE ÉQUIPE RENFORCÉE ET MOBILISÉE

- ▶ **Regroupement** de l'équipe modérateurs et de l'équipe administrateurs en une seule et même équipe HAL-CNRS et **arrivée de deux nouvelles collègues** : répartition des missions communes (modération, administration, assistance...) et définition des nouvelles missions à assurer (animation, accompagnement, formation, etc.).
- ▶ **Articulation des missions** de l'équipe autour de **deux objectifs** :
  - faire du portail institutionnel la vitrine de la production scientifique du CNRS,
  - soutenir l'objectif de 100 % de dépôts en texte intégral inscrit dans la feuille de route science ouverte du CNRS.

### FOCUS SUR... LES COLLECTIONS

**25 nouvelles collections d'unités CNRS** ont été créées en 2022 dont 4 créées, suivies et appuyées par l'Inist : Laboratoire de biologie physico-chimique des protéines membranaires (UMR7099 LBPC-PM) ; Institut des neurosciences Paris Saint-Pères (UMR8003 SPPIN) ; Institut de recherche en informatique fondamentale (UMR8243 IRIF) ; Institut de Biochimie et Génétique Cellulaires (UMR5095 IBGC).

Plus de **71 %** des UMR ont une collection dans HAL, reliée à HAL-CNRS.



Modération et mise en ligne de **68 040 dépôts**



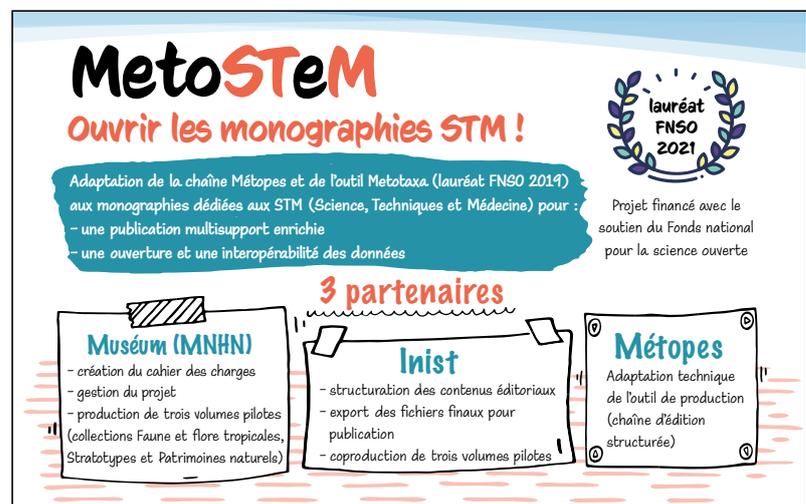
# ÉDITION NUMÉRIQUE

L'activité, l'équipe et les compétences ont été mises en lumière pour positionner l'Inist en tant que **soutien et acteur de l'édition scientifique française** et de la voie diamant.

## LANCEMENT DU PROJET METOSTEM

Lauréat du Fonds national pour la science ouverte 2021, ce projet réunit trois partenaires : **Muséum national d'histoire naturelle (MNHN), Métopes et Inist.**

L'objectif de MetoSTeM est d'adapter la chaîne Métopes et l'outil Metotaxa aux monographies dédiées aux STM (Science, Techniques et Médecine). Cette adaptation permettra une publication multisupport enrichie ainsi qu'une ouverture et une interopérabilité des données.



Séverine Koehl, Sylvie Lacour, Marie Berjon, Laurence Bénichou, Claire Remoué, et al.. **Metostem : ouvrir les monographies STM !. 10<sup>e</sup> journée du réseau Médici : réception et usages de la production scientifique**, Sep 2022, Toulouse, France. {hal-04003901}

## LA CO-ÉDITION AVEC LE MNHN CONFIRMÉE ET RENFORCÉE

Le **secrétariat de rédaction et d'édition** se poursuit dans le cadre de la revue *Anthropozoologica*.

**Trois collections d'ouvrages**, dont les membres de l'équipe sont éditeurs techniques pour les versions numériques, ont été diffusées sur SciencePress et OpenEdition.

## LE BADGE SPOC QUERO POUR L'ÉQUIPE DE L'INIST

Suite à la formation **Spoc Quéro, Qualité éditoriale référencement et outils**, l'équipe a obtenu un badge attestant de sa réussite. Ce dispositif de formation a pour but de favoriser la qualité éditoriale des publications.

Faciliter l'accès  
aux données, leur  
interopérabilité ainsi que  
leur réutilisation.



## OPIDOR

- ▶ **Des outils et services pour faciliter la gestion et la valorisation des données de recherche**

## DORANUM

- ▶ **Des formations adaptées pour la gestion des données de la recherche**

## EXPOSITION DES DONNÉES

- ▶ **Des données facilement accessibles en ligne au travers de sites web éditorialisés**

# OPIDOR

Le portail OPIDoR met à disposition un ensemble d'outils et de services pour faciliter la gestion et la valorisation des données de recherche en conformité avec les principes FAIR.

## L'INIST PORTE LE CENTRE DE RESSOURCES OPIDOR RECHERCHE DATA GOUV



Centre de ressources  
outils communs aux ateliers  
Recherche Data Gouv

Porté par l'Inist, il offre à l'ensemble de la communauté de l'enseignement supérieur et de la recherche des **outils et services nationaux pour la gestion des données**. Il :

- opère le service français d'**attribution des identifiants numériques pérennes de données (DOI)** à destination des établissements d'enseignement supérieur et de recherche ;
- met à disposition un outil généraliste de **création de plans de gestion de données** (plans de gestion de données activable par l'humain ou par une machine) ;
- accompagne les centres de référence thématiques dans le développement de la **déclinaison thématique des outils de création de plan de gestion de données**.



### PRÉSENTATION D'OPIDOR DANS L'ÉMISSION GRAND LABO DIFFUSÉE EN DIRECT SUR YOUTUBE

Mardi 15 février 2022, le thème de l'émission était « **Données de la recherche, quel partage ?** ».

À cette occasion, l'Inist a présenté OPIDoR notamment à travers les principes FAIR (faciles à trouver, accessibles, interopérables et réutilisables) et les plans de gestion de données.

L'émission est visible en replay sur YouTube.

## DMP OPIDOR : PROMOTION DU MODÈLE STRUCTURÉ

**OUTIL D'AIDE À LA CRÉATION EN LIGNE DE PLAN DE GESTION DE DONNÉES**, DMP OPIDoR est un des services du centre de ressources OPIDoR de Recherche Data Gouv.

Plusieurs actions ont été réalisées afin de le promouvoir :

- **Adoption du « machine actionnable DMP »** avec le formulaire personnalisé de Meso@LR.
- **Mise à disposition du format d'échange** RDA DMP Common Standard pour un DMP machine actionnable.
- **Mise à disposition de l'API.**
- **Interventions** (ex. ANR Tour, Journées RDA France, JCAD...).

# DATA CITE

**Membre du consortium international, l'Inist est agence d'attribution de DOI** (Digital Object Identifier) pour la France.

DataCite est une association à but non lucratif qui a pour objectif d'attribuer des identifiants numériques pérennes (DOI) aux jeux de données de la recherche dans leur définition la plus large : jeux de données, articles scientifiques, textes, logiciels...

## CONSORTIUM DATA CITE FRANCE

**PREMIÈRE ASSEMBLÉE GÉNÉRALE** le 1<sup>er</sup> décembre 2022.

**L'Inist a organisé la première assemblée générale du consortium français DataCite.**

Le CNRS, et par délégation l'Inist, est engagé auprès de DataCite pour piloter une adhésion consortiale des institutions françaises. Le consortium français permet de fédérer les adhésions des établissements français et francophones et de créer un lieu d'échanges pour toutes questions relatives aux identifiants uniques de type DOI.

Les sujets abordés lors de cette rencontre portaient sur les activités du consortium national, ses contributions au niveau international, une proposition d'organisation et d'animation du consortium national, l'élection des membres du CE.

45 personnes ont participé à cette assemblée.

## CONSORTIUM DATA CITE INTERNATIONAL

**PARTICIPATION AU GROUPE DE TRAVAIL SUR LES MÉTADONNÉES**

La version 4.5 du schéma des métadonnées s'est achevée avec notamment l'attribution d'un DOI aux « **Instruments** ».

**DÉVELOPPEMENT DE MÉTADORE**

En lien avec DataCite, l'Inist a travaillé sur le développement de l'**application MÉTADoRe** qui permet le signalement des données de la recherche dans BibCnrs.

# DORANUM

Plateforme de formation à distance, DoRANum permet de se former sur la gestion et le partage des données de la recherche.

## L'INIST PORTE LE CENTRE DE RESSOURCES DORANUM RECHERCHE DATA GOUV



Centre de ressources supports pédagogiques, e-formations  
Recherche Data Gov

L'Inist et le GIS Urfist mobilisent leurs compétences et la plateforme DoRANum sera enrichie par des supports de formation ou des e-formations travaillés avec les ateliers de la donnée en fonction des formations qu'ils dispensent sur leur site.

DoRANum capitalisera ainsi des supports de formation génériques et réutilisables par les ateliers de la donnée qui forment aux questions relatives aux données de la recherche.

Pour compléter les formations de proximité, des e-formations faciliteront l'auto-formation des équipes de recherche. Les ingénieurs pédagogiques de DoRANum pourront également apporter un soutien aux centres de référence thématiques.



## DES NOUVELLES RESSOURCES

En partenariat avec le GIS Réseau Urfist.

La plateforme a été enrichie suite à la réalisation de **12 nouvelles ressources pédagogiques génériques** et **2 nouvelles ressources disciplinaires** (projet DoRANum disciplinaire).

## OBERRED, UN ÉCOSYSTÈME DÉDIÉ AUX DONNÉES DE LA RECHERCHE

Le **projet européen Oberred**, auquel l'Inist participe, est arrivé à son terme après 36 mois de travail.

Réunissant 6 partenaires européens, Oberred est un projet Erasmus+ financé par la Commission européenne et piloté par l'université Côte d'Azur avec pour objectif de concevoir un écosystème pour reconnaître des compétences liées à la gestion et au partage des données de la recherche au travers d'open badges.

**Trois types de livrables ont été réalisés : des MOOC, un référentiel de compétences et les Open Badges.**

Bien qu'Oberred soit ouvert à tous, les Open Badges s'adressent principalement aux « facilitateurs » autrement dit tous les professionnels de l'IST qui travaillent à proximité des équipes de recherche comme les documentalistes ou les data steward, qui sont amenés à organiser des formations.

**Envie de tester ses connaissances et d'asseoir ses compétences avec un Open Badge reconnu ?  
Rendez-vous sur <https://oberred.eu>**

## CALLISTO, UNE PLATEFORME DE FORMATION EN MODE MOODLE

L'Inist met à disposition des tutoriels et webinaires sur Callisto.

Créée à l'initiative du ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation, Callisto est une plateforme de formation et d'autoformation à la documentation, aux métiers des bibliothèques et à l'information scientifique et technique.

Elle fédère les **ressources de différents partenaires** dont les Urfist, les CRFCB (Centre régional de formation aux carrières des bibliothèques) et maintenant l'Inist.

L'espace Inist de Callisto signale les ressources de formation sur les produits et service de l'Institut notamment autour d'Istex et de DMP OPIDoR. Il signale aussi les webinaires organisés par le service Formation.

Grâce à Callisto, les ressources de l'Inist sont visibles sur une **plateforme nationale mutualisée**. Callisto permet également aux formateurs des autres organismes d'intégrer les ressources de la plateforme dans leurs propres cours en ligne.

## GOPENDORE EN VERSION NUMÉRIQUE

LE JEU PÉDAGOGIQUE GOPENDORE sur la gestion des données de recherche a été développé en version numérique à la suite de son succès en version « cartes ».



Présenté sous forme de **Questions/Réponses** sur toutes les thématiques de DoRANum (plan de gestion, stockage des données, métadonnées...), il est composé de 3 parties de 7 questions. Une façon ludique de tester ses connaissances !

Une médaille est remise à la fin de chaque partie.

Envie de vous amuser tout en améliorant vos connaissances sur les sujets de la gestion des données de recherche ?

**À vous de jouer** : Le jeu GopenDoRe en ligne est accessible pour tous sur la plateforme DoRANum.



**Plus de 90 actions de formations** en présentiel ou à distance ont été dispensées en 2022 sur les sujets de l'ouverture et du partage des données et des réseaux sociaux scientifiques, pour **plus de 1500 participants**.



# EXPOSITION DES DONNÉES

Valoriser et partager les données de la recherche par la création de bibliothèques numériques avec OMEKA.

## PASSAGE D'OMEKA CLASSIC À OMEKA-S

L'Inist propose de rendre les données facilement accessibles en ligne au travers de sites web éditorialisés, à la demande des laboratoires, à l'aide de l'outil Omeka.

La **nouvelle version d'Omeka** est davantage tournée vers le web sémantique. En passant à Omeka-S, l'Inist propose ainsi de nouvelles fonctionnalités telles que la **technologie IIIF** pour les images en haute définition et la possibilité de réaliser des sites multilingues.

## CENHTOR

### UNE COLLABORATION AVEC LA MSH LORRAINE

L'Inist a collaboré pour la partie technique au **projet CENHTOR** (Centre de ressources numériques des humanités et des territoires) ; une plateforme de valorisation, d'exploitation et de diffusion des données de la recherche en Sciences humaines et sociales de la MSH Lorraine.

## UN SITE SUR LES REPRÉSENTATIONS ANIMALES

À la demande de l'Institut d'archéologie méditerranéenne Arkaia, l'Inist a créé une **plateforme collaborative** pour partager et valoriser des données de recherche sur les **représentations animales depuis la Préhistoire**.

La base de données, dénommée **Base Jean Clottes**, a été initialisée à partir de l'ouvrage *Bouquetins et Pyrénées, tome II : inventaire des représentations du Paléolithique pyrénéen*, publié en hommage à ce spécialiste de l'art préhistorique de renommée internationale, conservateur général du Patrimoine honoraire.



En 2022, l'Inist a travaillé sur la mise en oeuvre de **28 plateformes**, en collaboration avec **18 laboratoires** et **5 MSH**.



L'ensemble de ces plateformes hébergent **52 projets de recherche**.



Au total, **45 000 pages** ont été numérisées.

## UNE ENCYCLOPÉDIE EN LIGNE

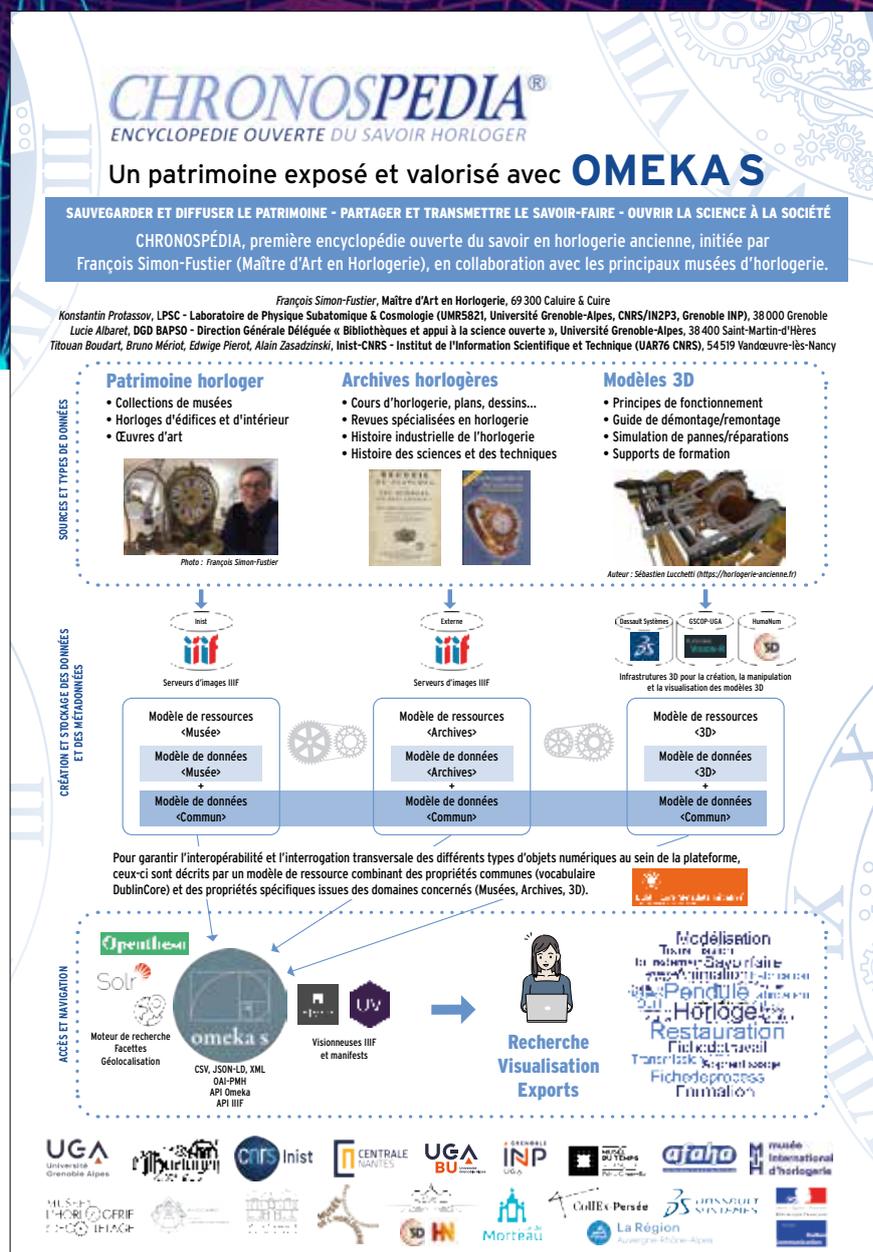
### UN PATRIMOINE EXPOSÉ ET VALORISÉ

L'Inist a mis en œuvre la base de données accueillant l'encyclopédie virtuelle « **Chronospedia** ».

La base rassemble des sources et types de données diverses (collections de musées, archives horlogères, modèles 3D...).

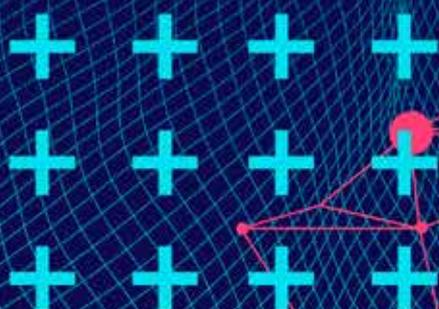
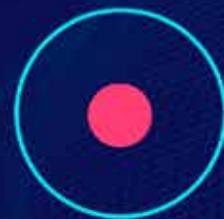
Le logiciel Omeka-S permet de rendre ces données accessibles en ligne et facilement consultables.

Initié par François Simon-Fustier, Maître d'Art en Horlogerie à la Croix-Rousse, ce projet réunit de nombreux partenaires. Il bénéficie d'un soutien financier de Collex-Persée 2022, du ministère de la Culture et de la Région Auvergne-Rhône-Alpes.



Timothy Collier, Silvia Falconeri, Isabelle Thiebau et Alain Zasadzinski,  
« **Sources pour une histoire juridique de la folie en situation coloniale** », Clio@Themis [En ligne], 23 | 2022, mis en ligne le 01 novembre 2022. URL : <http://journals.openedition.org/cliiothemis/2870> ; DOI : <https://doi.org/10.4000/cliiothemis.2870>

Répondre aux besoins de pilotage scientifique et d'extraction de connaissances en proposant une offre de services d'analyse et de fouille de textes.





## DES RESSOURCES POUR LE TEXT & DATA MINING (TDM)

- ▶ Istex, la plus importante source de données textuelles
- ▶ Loterre, la plateforme des ressources terminologiques

## INGÉNIERIE TERMINOLOGIQUE

- ▶ Des terminologies pour des travaux de classification, d'alignement, d'annotation, de fouille de textes...

## TEXT & DATA MINING (TDM)

- ▶ Des enrichissements adaptés aux besoins à l'aide d'outils automatiques

## ANALYSE ET VISUALISATION

- ▶ Une production scientifique suivie grâce à des indicateurs bibliométriques ou thématiques

## TRADUCTION

- ▶ Un appui linguistique pour la relecture ou la traduction de documents scientifiques

# DES RESSOURCES POUR

# LE TEXT & DATA MINING (TDM)

## ISTEX

LA PLUS IMPORTANTE SOURCE DE DONNÉES TEXTUELLES et ses outils intégrés mis à disposition de la communauté scientifique française pour l'analyse et la fouille de textes.

### PROJET D'INFRASTRUCTURE DE RECHERCHE

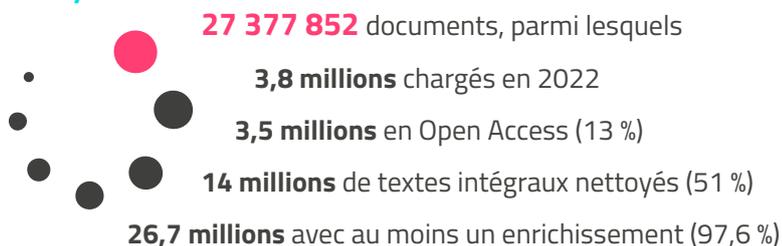
Istex fait partie des 108 infrastructures retenues dans la feuille de route nationale des Infrastructures de recherche éditée par le ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation (MESRI).

S'affichant comme une **bibliothèque scientifique numérique sans équivalent**, Istex a été présenté dans la catégorie projet d'infrastructure de recherche par son comité de pilotage constitué de représentants des différents partenaires (ABES, Consortium Couperin, Inist, France Université, Université de Lorraine).

Les principaux objectifs stratégiques exposés pour cette infrastructure sont :

- **ouvrir la collection** aux ressources nativement publiées en accès ouvert et poursuivre son alimentation grâce à une politique d'acquisition ;
- **faciliter la constitution** de corpus cohérents et enrichis, directement exploitables pour du TDM ;
- **promouvoir le développement** de services avancés avec la communauté des chercheurs en TAL (traitement automatique des langues) ;
- **offrir des services** d'exploration et d'exploitation de corpus accessibles à tous.

#### Istex, c'est :



#### ISTEX A 10 ANS

À l'occasion du **10<sup>e</sup> anniversaire d'Istex**, l'Inist et l'Abes ont écrit conjointement un article relatant le projet Istex devenu aujourd'hui le plus vaste réservoir d'archives scientifiques

Alexandra Petitjean-Monnin, Raluca Pierrot et Cécilia Fabry, « **ISTEX : de la plateforme de référence à l'infrastructure de recherche** », *Arabesques*, 107 | 2022, 20.



#### ISTEX FAIT PEAU NEUVE

Le **nouveau site Istex** a été lancé en mars avec un design plus moderne et une réorganisation des contenus.

S'inscrivant dans une démarche UX (User Experience), le site Istex évolue afin de mieux répondre aux besoins d'information et d'accompagnement sur toute la gamme de services proposés.

**DATA.ISTEX.FR** est un site, accessible à tous, qui expose des **jeux de données** Istex produits et/ou transformés à l'Inist selon les normes du web sémantique en les modélisant via une ontologie dédiée.

De **nouveaux corpus** y ont été publiés en 2022 :

**CORPUS D'ACTUALITÉ :**

- Abeilles
- Biomimétisme
- Organocatalyse asymétrique
- Végane & végétarien

- ClimatoRisques
- Blob

**CORPUS SPÉCIALISÉ :**

Mémoire-Neurosciences

Ces corpus scientifiques bénéficient des nouvelles ressources acquises dans le cadre des programmes CollEx-Persée, du Plan de soutien à l'édition scientifique française, et des acquisitions pérennes des Licences Nationales. Onze nouveaux **corpus éditeurs** sont ainsi venus enrichir le fonds Istex :

- Elsevier - années 2002-2012
- Brill – ebooks en religion antique
- Cambridge University Press – ebooks en sciences politiques
- DeGruyter – ebooks en langue française et en sciences de l'antiquité
- Casalini
- Proquest – Early European Books - bibliothèque nationale de Florence
- EDP Sciences - 2020-2021
- Proquest – Twentieth Century Religious Thought - Théologiens chrétiens du XX<sup>e</sup> siècle
- Proquest – Archives du Parlement britannique du XIX<sup>e</sup> siècle
- Cairn - revues en SHS
- Springer – ebooks en sciences politiques

**ISTEX.DL** est un site qui permet quant à lui d'**extraire un corpus de documents Istex**.

Afin de faciliter les extractions, et en réponse aux demandes d'utilisateurs, **des améliorations ont été réalisées** :

- «**TXT CLEANED**» : ce nouveau format permet de télécharger le **texte intégral nettoyé** (seul le contenu textuel des paragraphes est conservé).
- **PASSERELLE AVEC L'OUTIL CORTEXT** : conçue initialement à l'intention des chercheurs en SHS, la **plateforme CorText** est destinée à la fouille, l'analyse et l'exploitation de corpus textuels hétérogènes.
- **ASSISTANT DE CONSTRUCTION DE REQUÊTE** : un mode « **recherche assistée** » est disponible pour accompagner un utilisateur non averti.



**PLATEFORME D'EXPOSITION ET DE PARTAGE DE TERMINOLOGIES** scientifiques multidisciplinaires et multilingues.

## LOTERRE, NOUVELLE VERSION

La **version 3 de Loterre** a été lancée avec de **nouveaux services**.

Cette nouvelle version propose aux utilisateurs d'accéder à des services dédiés aux terminologies :

- **CONTRÔLER**, pour vérifier la validité d'un fichier SKOS/RDF-XML ;
- **TRANSFORMER**, pour convertir une terminologie dans différents formats, la corriger ou l'enrichir ;
- **ALIGNER**, pour aligner un fichier SKOS/RDF-XML avec une terminologie de Loterre ;
- **ANNOTER**, pour expérimenter l'annotation d'un texte avec une terminologie de Loterre.

Parmi les autres nouveautés, un plugin interroge Istex à partir du libellé des concepts. Cette fonctionnalité s'inscrit dans une démarche d'interopérabilité entre les différentes applications de l'Inist.

Un autre plugin affiche un aperçu de la page Wikipédia d'un concept, lorsqu'il a été aligné avec cette encyclopédie collaborative.

## RÉPONDRE À UN BESOIN SCIENTIFIQUE

Un **thésaurus de paléoclimatologie** a été publié sur Loterre.



Issu d'une réflexion avec des chercheurs du domaine, il a pour objectif d'offrir à la communauté scientifique un vocabulaire de référence bilingue (français, anglais), inexistant jusqu'alors.

Il a été réalisé en collaboration avec des chercheurs des laboratoires LSCE, Géosciences Montpellier et GEOPS (Géosciences Paris-Saclay) et une équipe de la DiBISO de l'Université Paris-Saclay.



# INGÉNIERIE TERMINOLOGIQUE

Faire entrer les terminologies dans l'ère du web de données.

## ASSISTER LA CONSTRUCTION DE TERMINOLOGIES AVEC TYDI

EN COLLABORATION AVEC L'INRAE ET L'UNIVERSITÉ PARIS-SACLAY.

L'Inist participe à la **phase 1 du projet TyDI** (Terminology Design Interface) ; projet qui vise à développer un **outil web collaboratif** de conception de vocabulaires structurés, à partir de listes de candidats termes issus d'extracteurs connectés en amont.

## ENRICHIR LE THÉSAURUS DE LA BIODIVERSITÉ

EN COLLABORATION AVEC L'ÉQUIPE PATRINAT (Centre d'expertise et de données sur le patrimoine naturel).

À la demande de l'équipe PatriNat, le but est d'enrichir le thésaurus de la biodiversité avec les données issues de la base de « **traits de vie** » qu'elle développe, l'un de ses objectifs étant par la suite d'utiliser le thésaurus en tant que référentiel de sémantisation d'une nouvelle base de connaissances.

## CO-ORGANISER UN ATELIER

EN COLLABORATION AVEC LE PNDB (Pôle national de données de biodiversité, infrastructure de recherche du MESR et centre de référence thématique de Recherche Data Gouv).

Les premiers jalons ont été posés afin de répondre au Conseil scientifique du PNCB qui souhaite organiser un atelier « **Enrichissement du thésaurus biodiversité de l'Inist et usage du thésaurus** » au Muséum national d'histoire naturelle en juin 2023.

Le but est d'enrichir les métadonnées associées au catalogue de données du PNDB.

## RÉALISER UNE VERSION XML

EN COLLABORATION AVEC L'INSERM.

Le projet consiste à réaliser une version XML du **thésaurus MeSH** (Medical Subject Headings) bilingue (français, anglais), pour les millésimes 2020, 2021 et 2022.

## MISE À JOUR DE RÉFÉRENTIELS

L'Inist a contribué à la mise à jour du **référentiel des lieux des Pactols**, par élaboration et fourniture à Frantiq (Fédération et ressources sur l'Antiquité) de référentiels actualisés et alignés avec GeoNames, Wikidata, IdRef, et Insee, pour la France, l'Allemagne, l'Espagne, l'Italie, le Portugal, la Grèce et les Balkans.

# TEXT & DATA MINING (TDM)

Analyser ou enrichir les documents scientifiques à l'aide d'outils automatiques (web services).

## RAPIDO RENDRE ACCESSIBLES DES PUBLICATIONS SCIENTIFIQUES INDEXÉES ET LIÉES À DES DONNÉES CERTIFIÉES

### LANCEMENT DU PROJET RAPIDO

Son objectif : **enrichir et mettre en relation les publications diffusées par Persée avec des sources externes** -référentiels, données de recherche- afin de les inscrire dans un environnement scientifique étayé par des données ouvertes, au moyen d'un protocole automatisé.

Le dispositif envisagé s'appuie sur le savoir-faire de l'Inist dans le domaine du **traitement automatisé de corpus** ainsi que sur la consolidation d'un partenariat effectif entre Persée et l'Abes autour de l'enrichissement réciproque de leurs données via le référentiel IdRef.

La collaboration porte d'abord sur un corpus délimité de publications des Écoles françaises à l'étranger diffusées sur le portail de Persée.

Persée porte le projet RAPIDO **lauréat du 2<sup>e</sup> appel à projets du Fonds national pour la science ouverte** en faveur de la publication scientifique ouverte.

RAPIDO est le fruit d'un partenariat entre Persée, l'Inist, l'EfR, l'École française de Rome, l'Efa (École française d'Athènes) et l'Abes (Agence bibliographique de l'enseignement supérieur).



## DES SERVICES EN TDM

### LANCEMENT DE LA NOUVELLE OFFRE DE SERVICES

Plusieurs **web services** ont été mis à disposition sur le site Objectif TDM.

Une **offre de formations** accompagne ces services.

## TEXTMINE

### LANCEMENT DU GROUPE DE TRAVAIL DE L'ASSOCIATION EGC

Le but de ce groupe de travail est de réunir des chercheurs sur la thématique large de la fouille de textes. Ce groupe vise à **offrir une occasion de rencontres** pour les universitaires et les industriels, appartenant aux différentes communautés de l'intelligence artificielle, l'apprentissage automatique, le traitement automatique des langues, pour discuter des méthodes de fouille de textes au sens large et de leurs applications.



Cuxac P., « **L'IA et la fouille de textes à l'INIST : l'IA à portée de tous ?** », *Arabesques*, 107 | 2022, 18-19.

Song K., Lamirel J.C., Cuxac P., Chen Y. : **Tracking evolution of technology topics based on an improved DP-BMM approach based on topic labelling.** 21<sup>st</sup> Collnet meeting, 10-12 Nov 2022, Bangkok, Thailand.

Bonvallet V., Parmentier F., Bourguignon L., Clauss I., Gregorio S., **Le TDM pour tous grâce à des web services au sein de LODEX, outil libre de visualisation**, *Revue des Nouvelles Technologies de l'Information*, Extraction et Gestion des Connaissances, RNTI-E-38, p. 445-452, 2022.

# ANALYSE ET VISUALISATION

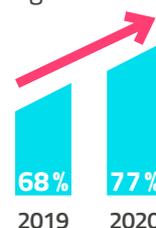
Les analyses bibliométriques permettent de suivre la production scientifique d'un laboratoire, d'un institut ou de toute autre structure, ou celle liée à une thématique scientifique et de dresser un panorama sur un sujet précis.



## LES PUBLICATIONS EN ACCÈS LIBRE QUANTIFIÉES

**LES PUBLICATIONS SCIENTIFIQUES EN ACCÈS LIBRE PRODUITES PAR LE CNRS** ont été quantifiées pour les années 2019 (à partir de la base Conditor) et 2020 (à partir des bases du Web of Science). Ce chiffre clé correspond à l'indicateur 18 du **Contrat d'objectifs et de performance 2019-2023 du CNRS** « Pourcentage des publications des unités dont le CNRS est tutelle publiées en accès ouvert ».

Sur les **47 382 publications CNRS** de 2020 issues des bases du Web of Science Core Collection de Clarivate Analytics, près de **77 % d'entre elles sont en accès libre**. Un chiffre en évolution puisque la part d'open access était d'environ 68 % pour les publications CNRS produites en 2019.



Évolution du nombre de publications CNRS en accès libre



Utilisé pour la mise à disposition des études bibliométriques, l'outil Lodex produit par l'Inist permet de **publier des données structurées** pour les présenter sous forme de graphiques.

## ÉTUDES BIBLIOMÉTRIQUES

L'Inist est régulièrement sollicité pour **caractériser l'activité scientifique** d'une structure ou d'une thématique scientifique. Parmi les études réalisées en 2022 :

- **réalisation d'indicateurs bibliométriques** à partir des résultats d'une expertise sur les impacts acoustiques des projets éoliens offshore sur la faune marine (MPES) ;
- **suivi de la production scientifique des unités de recherche de plusieurs instituts du CNRS** : INSU - INS2I - INSB - IN2P3 ;
- **suivi de la production scientifique de plusieurs unités de recherche du CNRS** : Chrono-environnement - CRIOBE, EDYTEM, RENSEE, SETE ;
- **étude des collaborations entre le CNRS et les CHU de France** pour la période 2016 -2020 (GT Santé du CNRS)
- **suivi de la production scientifique du Grand-Est dans les domaines de la pollution de l'air, du climat, et de l'énergie** pour la période 2007-2020 (ATMO Grand Est) ;
- **suivi de la production scientifique sur la thématique des côtes à falaises / rocheuses** 1960 à de nos jours (LETG) ;
- **suivi de la production scientifique française dans les domaines de la paléoclimatologie et du paléoenvironnement** pour la période 2000-2020 (LSCE - IGE) ;
- **suivi de la production scientifique française dans le domaine de la mémoire pour la période 2010-2020** (GDR Mémoire).

**DES OUTILS** pratiques développés par l'Inist et utilisés pour les études bibliométriques :



- **Webservices** : enrichissement par DOI (Unpaywall, Crossref) ; classification Pascal/Francis ; détection d'entités géographiques ; association d'un terme au vocabulaire Pays et Subdivision ; attribution de RNSR, des services de mapping (éditeurs/sources/types de documents/instituts CNRS).
- **Annotation des documents par détection** dans plusieurs champs des formes associées à chacun des concepts présents dans les thésaurus « Changement climatique » et « Pays et Subdivisions » développés à l'Inist et exposé sur sa plateforme terminologique ouverte Loterre.

## RAPPORT DU GIEC

**MISE EN ÉVIDENCE DE LA PLACE DU CNRS** dans les références bibliographiques citées dans le **6<sup>e</sup> rapport du GIEC** (DGDS - Direction générale déléguée à la science).

Sur les **50 757 publications** étudiées, il ressort que 8 % d'entre elles ont au moins un auteur français et que 72 % de ces auteurs français sont affiliés au CNRS.

Ainsi, **le CNRS se place 2<sup>e</sup> parmi les organismes mondiaux avec 3015 publications**, derrière l'Université de Californie qui compte 3212 publications.

## REPÉRAGE D'ADRESSES

Comme chaque année, l'Inist a participé à la **campagne de repérage d'adresses**, en partenariat avec le SAP2S (Service d'appui à la politique et à la prospective scientifiques), service de la Direction d'appui aux partenariats publics (DAPP) du CNRS.

L'Inist contribue donc au repérage des unités de recherche CNRS à partir des adresses enregistrées dans le Web of Science (WoS).

Véritable **outil d'aide à la décision**, les données issues de ce travail de repérage sont ensuite utilisées dans le rapport

d'activité annuel du CNRS ainsi que dans le rapport de données chiffrées et indicateurs.

Les résultats sont également utilisés par l'Observatoire des sciences et techniques (département du HCÉRES – Haut Conseil de l'évaluation de la recherche et de l'enseignement supérieur) pour la production des indicateurs de performance de la LOLF.

Les données peuvent par ailleurs servir dans le cadre d'études bibliométriques et de tableaux de bord réalisés pour les instituts du CNRS.

## TRADUCTION

Traduction, relecture de documents scientifiques et de supports de présentation : rendre la recherche accessible au plus grand nombre.



### LA SÉLECTION PAR L'ANR DU PROJET MATOS (Machine Translation for Open Science)

Le projet MaTOS a pour objectif de **fluidifier la circulation et la diffusion des connaissances scientifiques par une traduction automatique améliorée**.

Il vise à créer un **outil expérimental de traduction automatique** autour de la science, un domaine où cette technologie se heurte parfois à des difficultés.

MaTOS réunit **quatre partenaires** : LISN (Laboratoire interdisciplinaire des sciences du numérique) de l'Université Paris Saclay, CLILLAC (Centre de linguistique inter-langues, de lexicologie, de linguistique anglaise et de corpus-atelier), Inria et Inist.

### QUELQUES EXEMPLES DE TRAVAUX DE TRADUCTION EN ANGLAIS :

- **Rapport d'activité du CNRS** pour l'année 2021 notamment présenté par Antoine Petit lors d'une réunion à l'ONU ;
- **Critères d'évaluation** des sections du CoNRS ;
- **Contrat d'objectifs et de performance** du CNRS pour la période 2019-2023 ;
- **Rapport d'auto-évaluation 2021** notamment sur l'implication du CNRS dans le rapport de la COP21 (pour l'audit HCÉRES du CNRS).



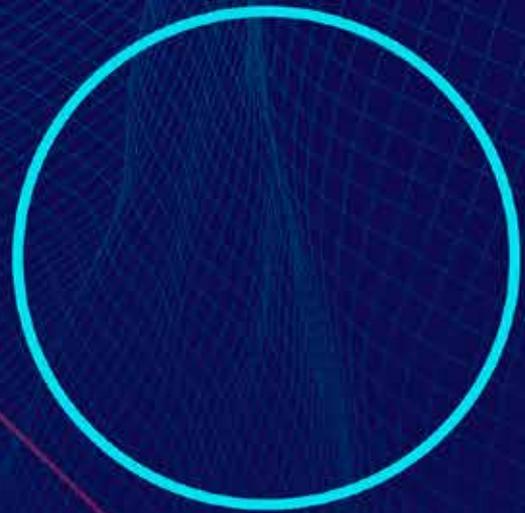
### LA TRADUCTION AU SERVICE DE LA SCIENCE OUVERTE

L'Inist a traduit plusieurs textes fondamentaux du **projet européen EOSC** (European Open Science Cloud) en collaboration avec Suzanne Dumouchel, représentante française dans l'association EOSC.

Le projet vise à offrir aux scientifiques européens, aux professionnels des sciences et de la technologie, un espace virtuel européen de gestion des données de recherche.

Quelque **40 000 mots** ont été traduits dans le sens anglais-français, et **20 500** dans le sens allemand-français ; un travail qui a pour objectif de rendre davantage accessibles les informations et de sensibiliser les acteurs de la recherche au projet EOSC.

L'Inist,  
c'est aussi...



# COMMUNICATION

## DES POSTERS

### PLUSIEURS POSTERS ONT ÉTÉ CONÇUS

et présentés lors de diverses rencontres :

- Le TDM pour tous grâce à des web services au sein de LODEX, outil libre de visualisation.
- ezCOUNTER : déléguez à ezMESURE la collecte des données d'usage COUNTER 5 de vos abonnements et économisez-vous un temps précieux !
- Nouvelle version de DMP OPIDoR : vers un DMP machine actionnable.
- MetoSTeM : ouvrir les monographies STM !
- Istex, 10 ans.
- Chronospedia, un patrimoine exposé et valorisé avec Omeka S.
- Inist-CNRS : de l'IST à la Science Ouverte.

## DES VIDÉOS « MÉTIERS »

### AFIN D'INCARNER LES MÉTIERS ET LES COMPÉTENCES

de l'Inist, quatre vidéos/ interviews ont été réalisées autour des métiers suivants :

- traducteur,
- chargé de négociations et d'acquisitions documentaires,
- ingénieur pédagogique,
- modérateur HAL.

Les vidéos sont visibles sur la **chaîne Youtube « Inist-CNRS »**.



4574 abonnés Twitter Inist

# VEILLE

## UNE VEILLE EN CONTINU AVEC LALIST

À TRAVERS SON SERVICE LALIST, véritable base de connaissances, l'Inist accompagne le monde de la recherche et les professionnels du domaine en mettant à disposition une mine d'informations actualisées.



1150 billets de veille publiés en 2022



1078 abonnés aux newsletters

## OUVRIR LA SCIENCE

L'Inist poursuit son **travail de veille pour le site *Ouvrir la science***, piloté par le Comité pour la science ouverte, en collaboration avec le ministère de l'Enseignement supérieur et de la Recherche pour la mise en ligne de contenus.

# WEB

## REFONTE COMPLÈTE DU SITE ISTEX

NOUVEAU DESIGN, NOUVELLES FONCTIONNALITÉS... L'Inist a mis en œuvre la **réalisation du nouveau site Istex**. S'inscrivant dans une démarche UX (User Experience), le site Istex a évolué afin de mieux répondre aux besoins d'information et d'accompagnement sur toute la gamme de services proposés.



# SECRETARIAT GÉNÉRAL



Au 31 décembre, l'Inist comptait **152 agents**

- ▶ **16** départs de l'unité
- ▶ **14** recrutements en concours
- ▶ **21** recrutements et contrats de courte durée établis

**BUDGET EXÉCUTÉ : 6 millions d'euros**  
dont environ 4,6 millions pour les acquisitions (BibCnrs)

Budget Acquisitions

## AUDIT PAR LE SERVICE FINANCIER ET COMPTABLE DU CNRS

Il résulte de l'audit une appréciation globale du risque de niveau modéré, qui démontre une appropriation ou une mise en œuvre satisfaisante des procédures de gestion comptable et financière, par l'unité.

# SYSTÈMES D'INFORMATION

## ÉTUDE DU DÉMÉNAGEMENT VERS LE DCML

L'Université de Lorraine, le CHRU de Nancy et la Métropole du Grand Nancy se sont engagés dans un projet de **Datacenter mutualisé lorrain (DCML)** dans le but de disposer d'un niveau de service renforcé et adapté aux besoins d'hébergement de leurs infrastructures. Le CNRS a intégré le projet en tant que partenaire opérationnel.

L'Inist a ainsi préparé le déménagement de ses données dans ce datacenter implanté sur le site Jean Zay, à côté de l'Institut.



### DES CAPACITÉS DE CALCUL EN HAUSSE

**6 serveurs physiques** ont été remplacés dans l'environnement virtualisé et des clusters DMZ, services internes, Windows et applicatif ont été réorganisés. Ces changements ont permis d'accroître de **25 % les capacités de calcul** pour atteindre **2,6 THz**.

## EN CHIFFRES...

**ENVIRONNEMENT VIRTUALISÉ** 28 serveurs physiques (ESXi)  
868 cœurs  
700 machines virtuelles

**STOCKAGE PRIMAIRE** Capacité brute **1000 To**  
répartie sur **4** baies

# LES RENCONTRES

## L'INIST EST INTERVENU AU COURS DE CONFÉRENCES ET FORMATIONS...

### JANVIER

**13.01 - Comité scientifique de DoRANum disciplinaire** - première réunion, **Claire Sowinski, Pierrette Paillassard, Lynda Kehli**

#### 20.01 - C@fé Renatis

Accompagner à la rédaction de PGD : retour d'expérience sur 3 approches pédagogiques, **Claire Sowinski**

**24-28.01 - 22<sup>e</sup> conférence sur l'extraction et la gestion des connaissances (EGC)**, Blois :

- Atelier TextMine, **Pascal Cuxac** ;
- Le TDM pour tous grâce à des web services au sein de LODEX, outil libre de visualisation (démonstration et poster), **Valérie Bonvallot**

### FÉVRIER

**08.02 - Atelier « ISTEEX pour la fouille de textes »**, Enssib, Villeurbanne, **Sabine Barreaux, Camille de Salabert**

### JUIN

**02.06 - Atelier pratique sur l'usage des ressources ISTEEX**, Toulouse, **Sabine Barreaux, Pascale Viot**

**09-10.06 - Journées CasuHAL** - Afficher une liste de publications dans WordPress : les plugins HAL et MonLaboratoire. Retours utilisateurs, **Théo Rousselle**

**21-22.06 - Journées de l'ABES**, Montpellier :

- ezCOUNTER : pas de SUSHI, on s'occupe de tout ! (poster), **Xavier Launois**
- Nouvelle version de DMP OPIDoR (poster), **Cécilia Fabry**
- Le TDM pour tous grâce à des Web Services dans LODEX (poster), **Valérie Bonvallot**

**30.06 - Atelier sur les bonnes pratiques en matière de stockage**, Colloque science ouverte, Université de Lorraine, **Yvette Lafosse**

### AOÛT

**28.08-02.09 - 62<sup>e</sup> édition du Groupe d'études en chimie organique (GECO)** - démonstrations et présentations, **Cécilia Fabry, Laurent Schmitt**

### SEPTEMBRE

**14-16.09 - Journées Médiçi**, Toulouse :

- MetoSTeM : ouvrir les monographies STM ! (poster), **Séverine Koehl** ;
- ReadMetrics : des outils et des services pour générer, récolter, exploiter et diffuser des données d'usage, **Yannick Schurter**

**27.09 - C@fé Renatis**

Les outils de formation en IST, **Claire Sowinski**

**28-30.09 - 51<sup>e</sup> congrès de l'ADBU**, Effets de la science ouverte sur les organisations, Caen :

- Inist-CNRS : de l'IST à la Science Ouverte (poster), **Claire François** ;
- Présentation des services, **Cécilia Fabry, Léa Antoine**

### OCTOBRE

**03-05.10 - ANF - TDM 2022 - Exploration documentaire et extraction d'information**, Paris :

- La constitution et l'exposition d'un corpus documentaire à partir des ressources IStex, **Sabine Barreaux, Camille de Salabert**
- Des web services dédiés de la fouille de textes, **Valérie Bonvallot, Justine Révol**

**21.10 - Journée de lancement du groupe de travail TextMine**, Paris, **Pascal Cuxac**

**24-28.10 - Open Access Week**, des publications et des données libres comme l'air, Paris :

- Panorama TDM, **Pascal Cuxac, Fabienne Kettani**
- Construction d'un corpus spécialisé à partir des ressources d'ISTEX, **Sabine Barreaux, Pascale Viot**

### NOVEMBRE

**14.11 - Cérémonie Talents CNRS NANCY 2022** - Remise du cristal collectif du CNRS à l'équipe **Click & Read**

**17.11 - Forum JobLab**, Nancy, **Cécilia Fabry**

**22.11-01.12 - 5<sup>e</sup> éd. des Journées RDA (Research Data Alliance)** :

- Les Plans de Gestion des Données – Bonnes pratiques de rédaction et outils (atelier), **Valérie Gadrat, Coralie Wysoczynski** ;
- Les Plans de Gestion des Données - L'outil DMP OPIDoR : modèle de plan de gestion de données structuré (atelier), **Yvette Lafosse, Laurent Rassinoux**

**28.11-01.12 - IncrLearn Workshop @ICDM**, Orlando (USA), **Pascal Cuxac**

**29.11 - Formation Enssib « La fouille de textes par l'exemple : du corpus à la représentation des résultats en passant par les outils »**, Villeurbanne, **Camille de Salabert et Pascale Viot**

### DÉCEMBRE

**01.12 - Première Assemblée Générale du Consortium national DataCite**, Inist, **Claire François, Paolo Lai**

**08.12 - Journée des doctorants de l'Atilf**, Nancy - IStex pour les linguistes, une ressource pour deux usages, **Mathilde Huguin**

**08-09.12 - Journées Omeka**, Grenoble, « Chronopedia, un patrimoine exposé en valorisé avec Omeka S » (poster), **Alain Zasadzinski, Edwige Pierot**

Ce rapport d'activité a été réalisé par le service Communication de l'Inist, en lien avec l'ensemble des services de l'Institut.

Conception graphique : Alexandra Petitjean-Monnin





**Inist - Institut de l'information scientifique et technique**

2, rue Jean-Zay - CS 10310  
54519 Vandœuvre-lès-Nancy

Tél. : +33(0)3 83 50 46 00

Fax : +33(0)3 83 50 46 50

[www.inist.fr](http://www.inist.fr)

