Inist

RAPPORT D'ACTIVITÉ 2020





Sommaire

p. 4 - ÉDITO

p. 5 - TROIS AXES FORTS

p. 6 - LES FAITS MARQUANTS

PAGE 8

ACCÉDER À L'INFORMATION SCIENTIFIQUE

p. 9 - Pour être au plus près des besoins : NÉGOCIATIONS ET MESURE DES USAGES

p. 10 - Pour accéder aux publications : PORTAILS & PLATEFORMES

p. 12 - Pour accompagner la mise en ligne de documents : ÉDITION NUMÉRIQUE

p. 14 - Pour accéder aux collections des laboratoires : HAL CNRS ET MODÉRATION HAL

PAGE 15

VALORISER LES DONNÉES DE LA RECHERCHE

p. 16 - À votre rencontre : OPIDOR TOUR ET FORMATIONS

p. 18 - Pour développer les pratiques FAIR : PARTAGE DES DONNÉES - OPIDOR

p. 20 - Pour tout savoir sur les données FAIR : FORMATIONS - DORANUM

p. 22 - Pour rendre visibles les données : EXPOSITION DE DONNÉES

PAGE 24

ANALYSER ET FOUILLER L'INFORMATION SCIENTIFIQUE

p. 25 - Pour des données enrichies et faciles à trouver : TEXT AND DATA MINING

p. 26 - Pour faciliter la fouille de textes : TEXTES & CORPUS - ISTEX

p. 28 - Pour un traitement optimal des terminologies : INGÉNIERIE TERMINOLOGIQUE

p. 30 - Pour un suivi sur les publications : APPUI AU PILOTAGE SCIENTIFIQUE

p. 32 - Pour un appui linguistique : TRADUCTION

PAGE 33

L'INIST, C'EST AUSSI...

p. 34 - Pour le fonctionnement de l'Institut : SECRÉTARIAT GÉNÉRAL

p. 35 - Pour répondre aux besoins de sécurité : SYSTÈME D'INFORMATION

p. 36 - Pour promouvoir les projets et services : COMMUNICATION - WEB

p. 37 - Pour le suivi de l'actualité : VEILLE

p. 38 - POUR EN SAVOIR PLUS...

Édito

Les services de l'Inist ouverts malgré tout



En cette année 2020 si particulière, l'Inist a poursuivi le déploiement de son offre de service positionnée dans le cadre de la **feuille de route pour la science ouverte du CNRS**.

Comme pour toute structure, la crise sanitaire a eu un impact fort sur le fonctionnement de l'Inist tout au long de l'année 2020.

Les services de l'Inist étant déjà dématérialisés, ils sont restés accessibles H24 7j/7, nous avons même observé une **augmentation notoire des consultations** sur nos plateformes documentaires BibCnrs et ISTEX et avons participé au partage de documentation spécifique avec un corpus intitulé « Coronavirus SRAS MERS » et le thésaurus COVID-19.

Cette **continuité de service** est le fruit d'un effort particulier, dès le premier confinement, pour outiller l'ensemble du personnel placé en télétravail et mettre en place très rapidement les outils de communication indispensables aux échanges et interactions intra-équipes. Je salue ici la réactivité et le professionnalisme de chacun.

L'Inist s'est impliqué pleinement dans le plan pour la science ouverte du CNRS et apporte son soutien opérationnel aux actions portées par la Direction générale déléguée à la science (DGDS) ou la Direction des données ouvertes de la recherche (DDOR).

Ainsi, à titre d'exemple, l'Inist a quantifié les publications scientifiques en accès libre produites par le CNRS (un des indicateurs du contrat d'objectifs et de performance de l'établissement) ; l'équipe en charge du portail HAL CNRS s'est pleinement impliquée pour modérer rapidement l'afflux considérable de dépôts des publications pour une prise en compte dans le compte rendu annuel d'activité des chercheurs (« campagne CRAC ») ; enfin l'équipe qui administre le catalogue de services dédiés aux données de la recherche (CatOPIDoR) travaille en étroite collaboration avec des représentants de tous les instituts pour structurer un répertoire des services de données portés par le CNRS.

Ces exemples sont les premières mises en application du positionnement de l'Inist comme « **bras armé** » du CNRS dans la bataille pour la science ouverte, comme l'a redit Antoine PETIT au personnel le 20 août 2020.

Tout au long de l'année, les plateformes de l'Inist continuent d'évoluer pour **répondre encore mieux aux besoins des chercheurs**. Le portail BibCNRS, le service de téléchargement « ISTEX DL (ISTEX Download) » et le site data Istex ont fait peau neuve pour apporter des fonctionnalités plus avancées et une navigation toujours plus simple et ergonomique. Ces évolutions sont définies en étroite collaboration avec les utilisateurs selon une **démarche « UX design »** qui permet de mieux comprendre leurs usages et leur ressenti vis-à-vis des interfaces que nous proposons. Je tiens à remercier ici nos collègues qui nous consacrent temps et retours.

Cette année encore et malgré les difficultés liées à la pandémie, l'Inist a réussi à garder une **proximité avec ses utilisateurs** en organisant des rencontres dans différents cadres et avec de nombreuses adaptations de formules et de calendrier : Opidor Tour initié en 2019 et finalisé en 2020, actions de formations sur la gestion des données de recherche, animation de formations Urfist sur la constitution de corpus spécialisés à partir du réservoir ISTEX et formation sur le TDM co-organisée avec l'ENSSIB.

La très grande majorité des projets de l'Inist est réalisée en partenariat, ce qui place l'Inist au cœur d'un réseau national d'acteurs très riche.

Cette année, l'alimentation de la plateforme ISTEX s'est poursuivie, entre autres dans le cadre des acquisitions du GIS Collex-Persée sous la forme d'une collaboration rapprochée avec l'Abes. L'évolution de DMP OPIDoR vers un service « Machine Actionable » interopérable avec les différents services utiles dans la gestion des données de recherche profite du projet commun ANR Flash avec l'IFB (Institut français de bioinformatique) et d'interactions fréquentes avec l'ANR et le GENCI. La plateforme de formation DoRANum réalisée avec le GIS Urfist est alimentée, non seulement par les réalisations de l'Inist, mais également dans le cadre de projets tel le projet DATACC piloté par nos collègues des universités de Lyon 1 et de Grenoble Alpes. Un vocabulaire normé en paléoclimatologie, référence nécessaire pour documenter les jeux de données du domaine, est en cours de réalisation avec l'Université de Paris-Saclay. Enfin, le Comité pour la science ouverte (CoSO) s'appuie sur les collègues de l'Inist pour gérer et alimenter son site « Ouvrir la science ».

Les différentes réalisations présentées dans ce rapport d'activité sont les pièces d'un puzzle qui nous permet de dessiner l'Inist de demain, réflexion fondamentale pour consolider nos moyens et ancrer dans le temps cette dynamique.

Bonne lecture.

Claire FRANÇOIS
Directrice de l'Inist

Trois axes forts

Unité d'appui à la recherche (UAR), l'Inist déploie ses activités autour de trois axes forts « Accès à l'information scientifique », « Valorisation des données de la recherche », « Analyse et fouille de l'information scientifique », et d'un axe de services d'accompagnement.
Ce projet s'inscrit dans le contexte de politique institutionnelle et nationale de science ouverte.

Accéder à l'information scientifique

Pour s'appuyer sur les résultats de la recherche diffusés dans les publications internationales, faciliter l'accès à ces sources d'information ou encore constituer son propre corpus d'articles, l'Inist met à disposition des plateformes de ressources documentaires et des services associés.



Valoriser les données de la recherche

Pour faciliter la gestion, le partage, la diffusion ou l'archivage des données issues de travaux de recherche, l'Inist propose un bouquet de services et d'outils dédiés, ainsi qu'un accompagnement personnalisé incluant des modules d'autoformation.

Analyser & fouiller l'information scientifique

Pour répondre aux besoins de pilotage scientifique et d'extraction de connaissances, formulés par des structures institutionnelles, fonctionnelles ou des opérateurs de la recherche, l'Institut opère des services d'analyse et de fouille de corpus textuels scientifiques.

La science ouverte,
fil conducteur de ce rapport d'activité.

JANVIER

Conditor devient un service opérationnel

FÉVRIER

- La production du CNRS en libre accès quantifiée
- Accord EDP Sciences acta acustica
- Mise en ligne de l'archive de la BDSP

MARS

COVID-19 : Accès garantis à nos services et outils

COVID-19 : Généralisation du télétravail

AVRIL

L'usage des ressources électroniques grimpe en flèche

MAI

Nouvelle version data.istex : une navigation simple et ergonomique

JUIN

- Tableaux de bord trimestriels ezMESURE
- Les corpus ISTEX présentés à TALN en conférence virtuelle
- BibCnrs nouvelle version est en ligne !
- DoRANum publie un parcours interactif



JUILLET

Premier corpus d'actualités sur data.istex : « Coronavirus SRAS MERS »

AOÛT

Un thésaurus COVID-19 dans Loterre

SEPTEMBRE

Contribution à la mise à jour du modèle DMP Science Europe (grille d'évaluation des DMP)

OCTOBRE

- Co-organisation des Journées Omeka à Nancy
- Évaluation « FAIRitude » pour Loterre et TermSciences

NOVEMBRE

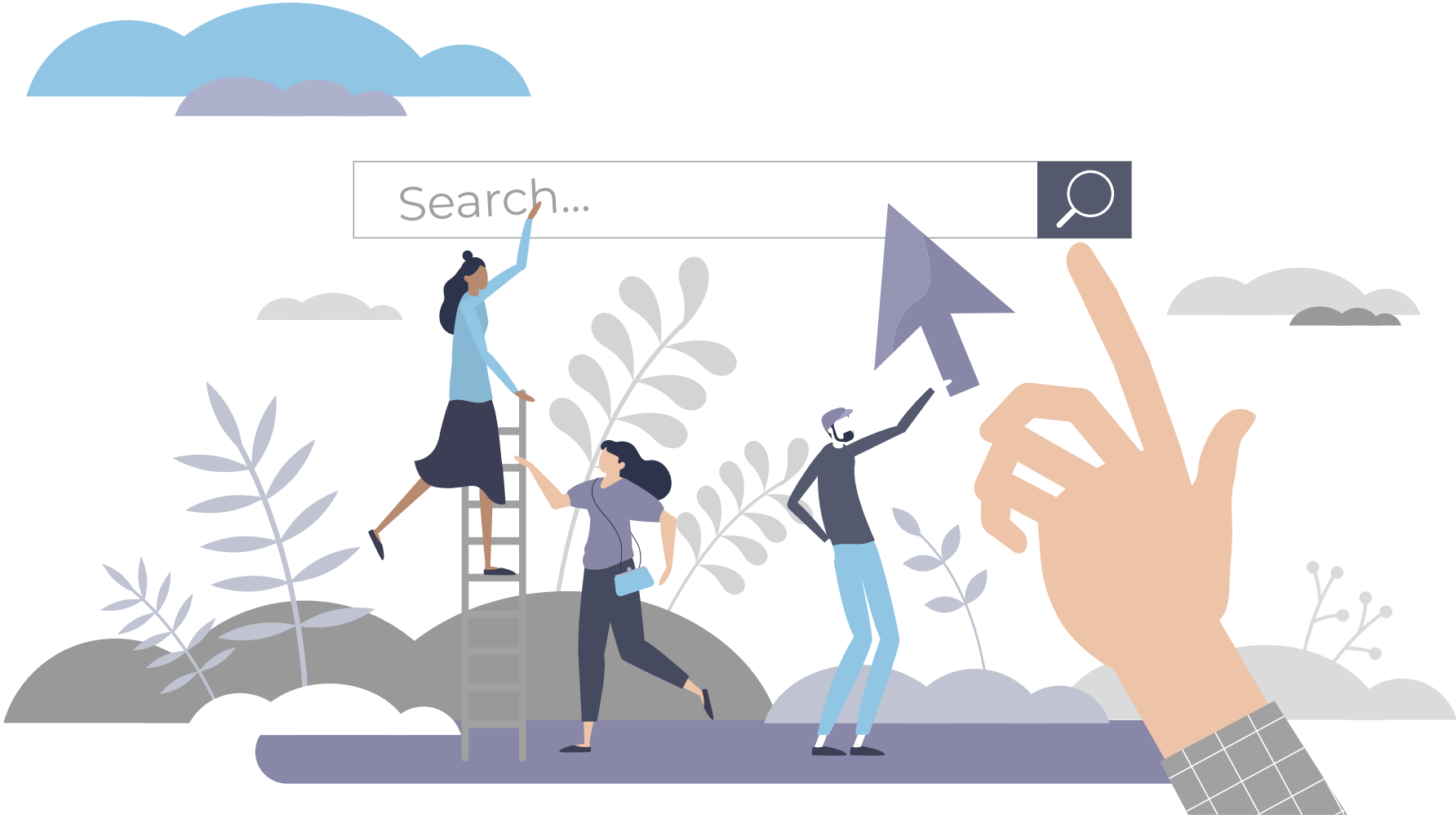
- ISTEX-DL, nouvelle version !
- Le service Publications participe à la nouvelle collection du MNHN
- Mise en ligne d'un nouveau site OMEKA *Revues Coloniales Européennes*

DÉCEMBRE

DMP OPIDoR évolue pour être exploitable par les machines



Accéder à l'information scientifique



Pour être au plus près des besoins : **Négociations et Mesure des usages**

Le service Négociations travaille au renouvellement et à la mise en place de nouveaux abonnements notamment pour le portail BibCnrs, les discussions portant sur les tarifs, les licences ou encore les conditions d'accès.
Le service Mesure des usages - ezMESURE intervient dans les orientations de la politique documentaire en identifiant les titres les plus consultés ainsi que les communautés utilisatrices.



Le fait majeur

Des **pics de consultation** BibCnrs ont été constatés pendant le confinement, les hausses allant jusqu'à 65 %. Sans surprise, les articles les plus consultés concernaient la COVID-19 et les coronavirus.
Durant cette période de travail à distance et d'effervescence dans la recherche en virologie et en médecine, BibCnrs s'est montré incontournable pour l'ensemble des domaines scientifiques.

En avril 2020, un accord a été signé avec EDP Sciences : il couvre le financement des frais de publication dans la revue *Acta Acustica* en open access. Il donne la possibilité aux chercheurs des laboratoires du CNRS en acoustique de publier jusqu'à 52 articles sous licence CC-BY-4.0 sur trois années ou jusqu'à épuisement du montant initialement investi.

Les autres évènements de l'année...

L'Inist élabore chaque année des **tableaux de bord** pour suivre les usages des ressources électroniques. Basés sur les données d'usage COUNTER mais également sur les données financières, ils sont maintenant déclinés **par institut** pour les ressources thématiques ou partagées par plusieurs communautés. Ces indicateurs sont établis a posteriori, quand une année complète de statistiques est disponible. En complément, des tableaux de bord sont envoyés de **façon automatique** tous les trois et six mois à nos correspondants IST dans les instituts afin de disposer de données sur la période en cours.

Début du travail sur l'import COUNTER 5 dans ezMESURE. L'objectif de cette fonctionnalité est de faciliter la collecte des données d'usage éditeur en permettant une automatisation du processus via l'usage du protocole SUSHI mis en œuvre dans COUNTER 5. Les utilisateurs saisissent leurs identifiants et l'application fait le reste. L'Inist a déjà renseigné 45 plateformes éditeurs.

Annonce ReadMETRICS : mise à disposition en Open Source d'une stack facilement déployable constituée d'ezPAARSE et d'ezMESURE avec appel à financement participatif permettant de renforcer les opportunités de collaboration au niveau européen.

Partenariats et collaborations

- **Our_research** : accord pour la mise en place d'un serveur miroir Unpaywall permettant notamment de rapatrier les informations d'enrichissement via le DOI entre autres pour l'Open Access et d'économiser des requêtes Crossref.

Plus de **50** négociations
34 millions d'articles consultés
68 établissements français contributeurs
470 plateformes déclarées dans Analogist
325 parseurs actifs
589 abonnés Twitter (**130** tweets)

Pour accéder aux publications : **Portails & plateformes**

L'Institut héberge et gère trois plateformes permettant l'accès à des documents scientifiques électroniques tant pour le CNRS que pour l'ensemble des établissements de l'enseignement supérieur et de la recherche (ESR) :

BibCnrs
Portail d'accès aux ressources électroniques du CNRS, il met à disposition des communautés de recherche du CNRS un ensemble négocié de plus de 19 000 revues, 18 000 ouvrages et de 40 bases de données multidisciplinaires, complété de ressources spécifiques à chaque institut du CNRS.

ISTEX
Réservoir d'archives de plus de 23 millions de documents accessibles à l'ensemble de l'ESR, la plateforme ISTEX est un complément aux abonnements courants du portail BibCnrs. Elle garantit un accès pérenne aux collections rétrospectives acquises.

PANIST
Plateforme d'accès pérenne, commanditée par le Consortium Couperin, elle offre aux établissements de l'ESR ayants droit la garantie de l'accès aux contenus souscrits dans les contrats indépendamment de l'évolution des marchés avec les éditeurs.

Le fait majeur

15 juin : lancement du nouveau site BibCnrs !
BibCnrs a connu et va connaître encore des évolutions afin d'être toujours au plus près des besoins des chercheurs.
En réponse aux attentes issues d'une démarche UX (expérience utilisateur), une nouvelle interface plus épurée a été créée pour simplifier la navigation et mettre d'avantage en évidence la recherche :

- **accès au texte intégral** facilité en un seul clic ;
- **nouvelle charte graphique** plus moderne et plus lisible ;
- possibilité pour les utilisateurs de **personnaliser leur espace** et d'épingler des ressources favorites pour les consulter plus facilement.

Les autres évènements de l'année...

ISTEX : initialement inscrite dans l'axe « Analyser et fouiller l'information scientifique » pour l'enrichissement et la fouille de texte, **l'offre de service ISTEX a été développée sur un second axe** « Accéder à l'information scientifique » pour le signalement des ressources dans la plateforme ISTEX. Ce déploiement a induit une phase de transfert de savoirs internes et une montée en compétences de l'équipe : contractualisation de la réception des données ou métadonnées éditeur et analyse détaillée des corpus livrés.

Démarrage en novembre du projet de bouton unifié Click & Read.
L'objectif de la nouvelle extension est d'unifier les trois programmes existants (BibCnrs, ISTEX et PANIST) pour n'avoir à en gérer plus qu'un seul afin de répondre à la problématique constante de simplification des accès et de gain de temps pour un accès plus rapide au texte intégral.



Partenariats et collaborations

- **Abes :** l'Inist intervient en tant qu'opérateur technique au côté de l'Abes en charge des achats en licences nationales dans le cadre :

  · du projet du **GIS CollEx-Persée** d'acquisitions de ressources ciblées à destination des établissements de l'ESR en garantissant à terme l'accès pérenne aux collections via la plateforme ISTEX ;

  · du **plan de soutien à l'édition scientifique française** avec la mise à disposition via la plateforme ISTEX fin 2020 des données de l'éditeur EDP Science 2017-2019.

- **Consortium Couperin :** l'Inist héberge et garantit l'accès à la plateforme PANIST commanditée par le Consortium Couperin.

Merci!

L'Inist remercie l'ensemble des participants à la démarche expérience utilisateurs qui ont contribué à faire évoluer et simplifier leur portail BibCnrs.

L'évolution de BibCnrs a permis la mise en avant des ressources libres via le pictogramme OA directement lors d'une recherche comme HAL, DOAJ, ArXiv, ainsi que dans l'implémentation de l'appel à Unpaywall au niveau de la recherche article.

Actions de communication

- **14 webinaires** pour présenter la nouvelle version de BibCnrs (391 participants)
- Diffusion d'un **sketchnote** pour le lancement de BibCnrs



BibCnrs :

6 356 656 consultations ezPAARSE dont :

· 2,9 millions d'articles scientifiques sous licence éditeur et 62 109 consultations d'articles issus de ressources libres ;

· 1,8 million de recherches via les bases de données

50 464 chercheurs ayants droit : 13 278 chercheurs CNRS répartis dans 1 537 unités + 37 186 chercheurs non CNRS

6 646 abonnés aux listes de diffusion

1 068 abonnés au compte Twitter

ISTEX : 666 380 consultations documentaires

PANIST : 3 702 consultations



Pour accompagner la mise en ligne de documents : **ÉDITION NUMÉRIQUE**

L'équipe est formée au stylage et à l'utilisation de la chaîne éditoriale Métopes qui a la particularité de gérer le format XML pour la diffusion multi-supports, assurant ainsi une pérennisation des contenus et une haute qualité des métadonnées associées. Cette nouvelle expertise a permis de monter de nouvelles collaborations autour de la valorisation de revues scientifiques, en complément de l'activité de secrétariat de rédaction.

Le fait majeur

L'Institut a collaboré avec le Muséum national d'histoire naturelle (MNHN) dans les trois premières parutions de la nouvelle collection « Natures en sociétés ».
En effet, fort de ses compétences en matière de stylage et dans l'utilisation de la chaîne éditoriale Métopes, le service Édition numérique a été sollicité par le Muséum pour sa nouvelle collection en numérique et en version papier.
L'Inist avait notamment en charge la préparation des manuscrits dans la chaîne Métopes, les contrôles biblio-typo-ortho-grammaticaux, le traitement des images quand nécessaire ou encore la cohérence des contenus.

Les autres évènements de l'année...



***Arabesques* en ligne :** collaboration avec l'Abes pour la mise en ligne de la revue *Arabesques* sur Prairial, portail des revues de l'Université Lyon 3. Avec l'objectif d'améliorer la diffusion en ligne de sa revue, l'Abes a fait appel à l'Inist pour sa connaissance de la chaîne éditoriale Métopes, afin de rendre disponible en ligne la collection complète de la revue sur Prairial. Portail de revues scientifiques, Prairial a obtenu en 2018 le statut de pépinière officielle d'OpenEdition et s'inscrit dans le paysage de la science ouverte.

Arrêt de IRevues : la plateforme est arrêtée depuis le 31 décembre. Les revues des éditeurs qui le souhaitaient ont été transférées vers d'autres plateformes.

L'équipe Édition numérique a participé au traitement de numéros de plusieurs revues en SHS qui ont été ensuite chargées sur les plateformes d'**OpenEdition**.

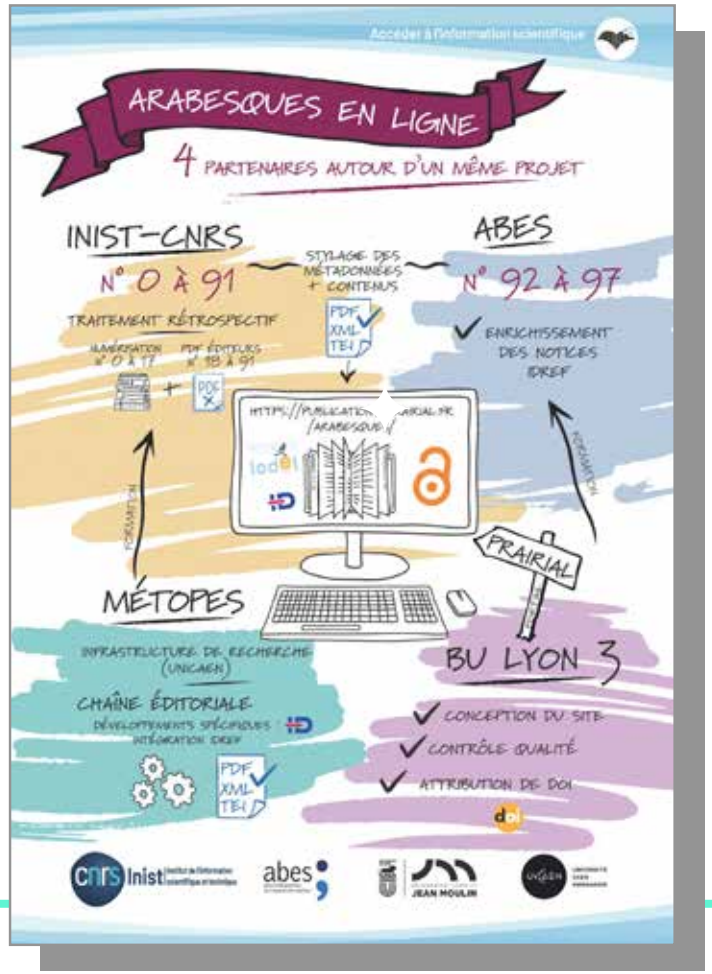
Le travail d'édition numérique contribue à mettre en ligne des documents accessibles gratuitement pour les chercheurs et l'ensemble des acteurs de l'enseignement supérieur et de la recherche.

Partenariats et collaborations

- **MNHN :** secrétariat de rédaction et d'édition réalisé à l'Inist pour le périodique scientifique *Anthropozoologica*.
  Collaboration dans la parution des trois premiers volumes de la nouvelle collection « Natures en sociétés ».
- **Abes :** participation dans le projet de mise en ligne de la revue *Arabesques*.
- **Open Edition Center :** collaboration dans l'édition de plusieurs revues et ouvrages.

Actions de communication

- ***Arabesques en ligne : une collaboration fructueuse***, article publié sur le site web Inist (lire p. 40).
- ***Le service Publications participe à la nouvelle collection du MNHN***, article publié sur le site web Inist (lire p. 45).
- Poster ***Arabesques en ligne – 4 partenaires autour d'un même projet***, préparé pour les Journées de l'Abes, 5 et 6 mai 2020 (manifestation annulée en raison de la crise sanitaire).



7 numéros d'*Arabesques* traités pour mise en ligne

3 ouvrages parus de la collection « Natures en sociétés » (MNHN)

60 numéros mis en ligne pour OpenEdition

19 articles édités d'*Anthropozoologica* (MNHN)



Pour accéder aux collections des laboratoires : **HAL CNRS et modération HAL**

L'Inist s'implique fortement dans le soutien de l'archive ouverte HAL en administrant le portail HAL CNRS dédié à la valorisation des productions scientifiques du CNRS et en assurant la modération de dépôts.

Le fait majeur

Campagne CRAC et impact sur la modération

Dans le cadre de sa politique en science ouverte, le CNRS a rendu obligatoire le dépôt des publications de ses chercheurs dans l'archive HAL pour une prise en compte dans les rapports annuels d'activité des chercheurs (CRAC).
Cette injonction a provoqué une augmentation considérable du nombre de dépôts dans HAL en novembre et décembre.
De nouvelles modalités de modération avec le CCSD ont été mises en place pour faire face à cet afflux de dépôts :

- Traitement allégé de la modération : nouvelles règles pour la modération à appliquer aux dépôts CRAC ;
- Traitement prioritaire des dépôts concernant le CRAC (année publication 2019-2020) ;
- Renforcement des équipes de modération.

Partenariats et collaborations

- **CCSD :** modération des dépôts dans HAL et soutien aux dépôts (conseil, incitation, aide).

Les autres évènements de l'année...

Création de collections dans le portail HAL CNRS :
Le travail d'administration du portail porte également sur l'ouverture des collections des laboratoires qui complèteront les quelques 750 laboratoires déjà reconnus dans le portail. Cette action fait suite à l'enquête diffusée par la DIST début 2020 auprès des laboratoires. L'objectif est d'intégrer le plus de laboratoires pour que le portail soit le plus exhaustif possible.

Ouverture de la **collection Inist** dans le portail HAL CNRS.

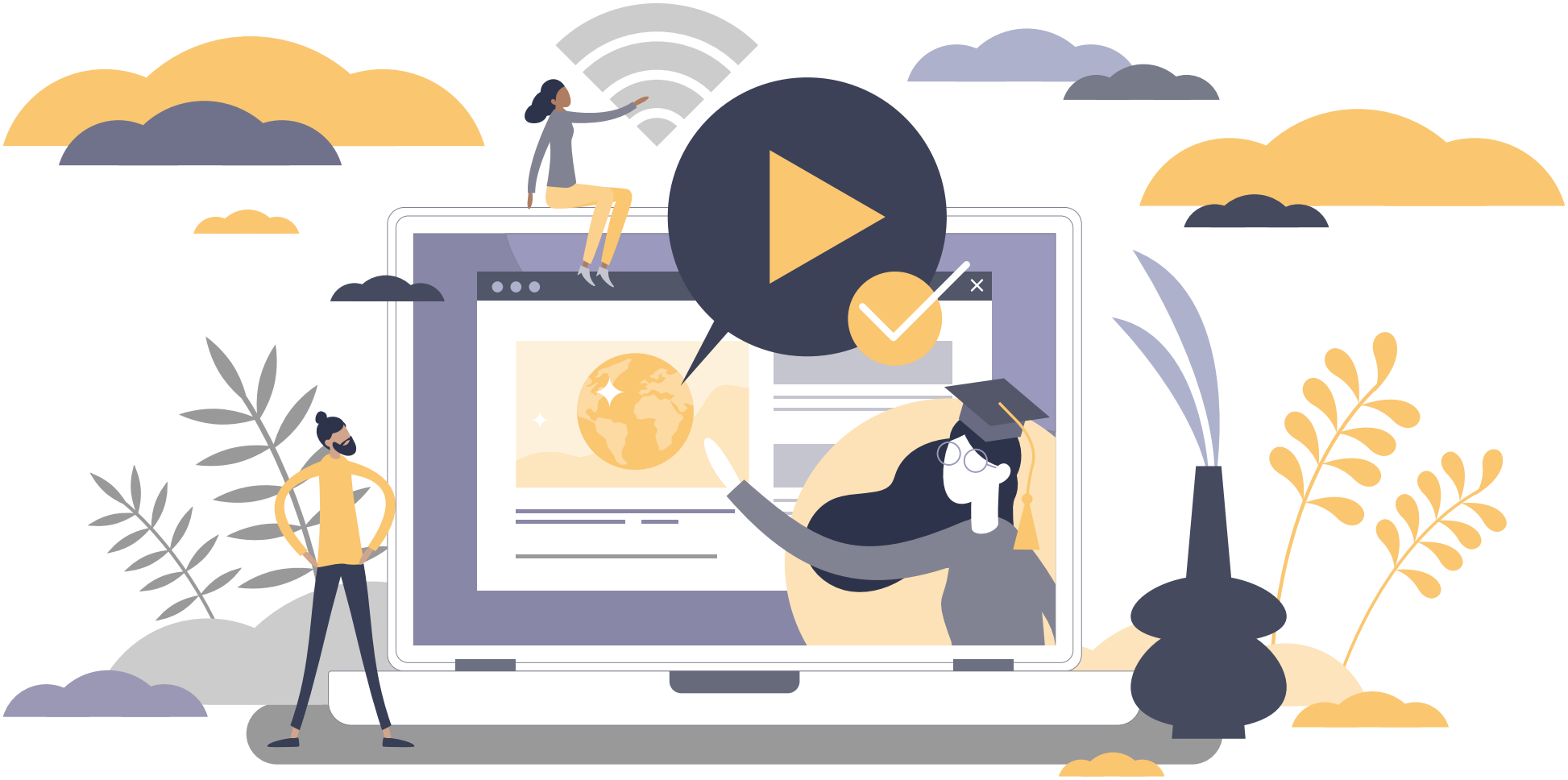
Portail HAL CNRS : 692 laboratoires (sur 1 137) y ont une collection

Modération dans HAL :
60 788 dépôts traités par l'Inist dont 6 158 affiliations corrigées
44 666 publications mises en ligne dans HAL
15 479 demandes de modifications ou conseils apportés aux déposants

CRAC : 500 dépôts supplémentaires journaliers

En administrant le portail HAL CNRS, l'Inist contribue à rendre accessible l'ensemble des publications des chercheuses et chercheurs du CNRS.

Valoriser les données de la recherche



À votre rencontre : **Opidor Tour et formations**

Lancé en 2019 par l'Inist avec les délégations régionales du CNRS, en lien avec les services de proximité des universités, l'Opidor Tour s'est poursuivi en 2020.

Des formations ont également été organisées, en présentiel et en distanciel.



Opidor Tour 2020
475 participants

Opidor Tour DR 19 - 17 janv.
Coralie WYSOCZYNSKI
Yvette LAFOSSE
Jean-Michel PARRET

Opidor Tour DR 17 - 3 mars
Yvette LAFOSSE
Françoise TISSERAND

Caen (visio)

Rennes

Rouen (visio)

Opidor Tour DR 19 - 16 janv.
Coralie WYSOCZYNSKI
Yvette LAFOSSE
Jean-Michel PARRET

Lille

Opidor Tour DR 18 - 10 mars
Coralie WYSOCZYNSKI
Laurent RASSINOUX

15

Gif-sur-Yvette

3 4 9 12 13 14

Opidor Tour DR 04 - 10 juin
Yvette LAFOSSE
Jean-Michel PARRET

Guyancourt

Opidor Tour DR 04 - 20 janv.
Pierrette PAILLASSARD
Yvette LAFOSSE
Marie-Christine JACQUEMOT PERBAL

2 5

Strasbourg

Opidor Tour DR 10 - 14 janv.
Pierrette PAILLASSARD
Claire SOWINSKI
Laurent RASSINOUX

Nantes

Opidor Tour MSH - 2 mars
Yvette LAFOSSE
Françoise TISSERAND

Villeurbanne

7 11 1

Opidor Tour DR 07 - 7 janv.
Coralie WYSOCZYNSKI
Yvette LAFOSSE
Françoise COSSERAT

Talence

Opidor Tour DR 15 - 12 mars
Coralie WYSOCZYNSKI
Anne BUSIN

Toulouse

8

Opidor Tour DR 14 - 5 mars
Yvette LAFOSSE
Paolo LAÏ

Montpellier

Opidor Tour DR 13 - 28 janv.
Coralie WYSOCZYNSKI
Yvette LAFOSSE
Bernard SAMPITÉ

Marseille

Opidor Tour DR 12 - 4 fév.
Claire SOWINSKI
Coralie WYSOCZYNSKI
Mohamed YAHIA

10

Nice

Opidor Tour DR 20 - 6 fév.
Claire SOWINSKI
Coralie WYSOCZYNSKI
Mohamed YAHIA

Nos interventions :

1. **Hackathon : comment améliorer le dépôt et le partage de données de recherche ?**
   50 participants - Grenoble – 14 janvier
   Jean-Michel PARRET, Françoise TISSERAND, Mohamed YAHIA, Loys BERTHELIER
2. **Formation - Atelier DMP - UL**
   11 participants - UL - Nancy, 23 janvier
   Claire SOWINSKI, Pierrette PAILLASSARD, Thomas JOUNEAU (UL)
3. **Formation - Science ouverte et gestion des données de la recherche**
   24 participants - ESPCI - Paris, 5 février
   Yvette LAFOSSE
4. **Journées « Préserver ses données numériques », TGIR Huma-Num**
   64 participants - Paris, 7 février
   Paolo LAÏ
5. **Rencontre Administrateurs des laboratoires de la DR Centre-Est**
   25 participants - Vandœuvre-lès-Nancy, 10 février
   Paolo LAÏ
6. **Journées IDCC Démonstration DMP OPIDoR (Produits de recherche)**
   250 participants - Dublin, 18 février
   Benjamin FAURE
7. **Atelier : « Le DMP ANR, comment le renseigner, l'outil DMP OPIDoR »**
   20 participants - Lyon (visio), 20 février
   Anne CIOLEK, Françoise COSSERAT, Jean-Michel PARRET
8. **Défi 13 INSU : workflow des données et DMP OPIDoR**
   44 participants - Toulouse (visio), 6 mars
   Anne CIOLEK, Jean-Michel PARRET
9. **Formation - Atelier de sensibilisation DoRANum et OPIDoR**
   38 participants - Paris BU 8 (visio), 20 mai
   Claire SOWINSKI, Yvette LAFOSSE
10. **Atelier DMP - Open Access Week Urfist Méditerranée**
    38 participants - Nice (visio), 15 juin
    Coralie WYSOCZYNSKI, Claire SOWINSKI
11. **Atelier DMP**
    31 participants - GRICAD Grenoble (visio), 7 juillet
    Paolo LAÏ, Yvette LAFOSSE
12. **« DMP : comment en améliorer le contenu ? » Open Access Week**
    19 participants - Paris 08 (visio), 20 octobre
    Yvette LAFOSSE, Claire SOWINSKI, Coralie WYSOCZYNSKI, Pierrette PAILLASSARD, Justine ANCELIN (Urfist Paris)
13. **« DMP : relecture et analyse du contenu d'un plan » Open Access Week**
    6 participants - Paris 08 (visio), 20 octobre
    Yvette LAFOSSE, Claire SOWINSKI, Coralie WYSOCZYNSKI, Pierrette PAILLASSARD, Justine ANCELIN (Urfist Paris)
14. **« Atelier DMP : relecture et analyse du contenu d'un plan »**
    8 participants - Paris 08 (visio), 3 novembre
    Yvette LAFOSSE, Claire SOWINSKI, Coralie WYSOCZYNSKI, Pierrette PAILLASSARD, Justine ANCELIN (Urfist Paris)
15. **Atelier DMP et DMP OPIDoR**
    53 participants - Université d'Artois (visio), 23 novembre
    Laurent RASSINOUX, Claire SOWINSKI

Webinaires organisés par le service formation-DoRANum (DoRANum, DMP OPIDoR et Réseaux sociaux)

- **Partager et valoriser ses recherches via les réseaux sociaux**
  77 participants - 22 septembre, 13 octobre, 17 novembre
  Valérie GADRAT, Pierrette PAILLASSARD
- **Présentation de DoRANum**
  245 participants - 9 juin, 1er et 6 octobre, 24 septembre, 24 novembre, 15 décembre
  Yvette LAFOSSE, Philippe FOURNERET, Claire SOWINSKI
- **Présentation de l'outil DMP OPIDoR**
  454 participants - 25 juin, 2 juillet, 17 septembre, 1er et 15 octobre, 5 et 19 novembre, 8 décembre
  Claire SOWINSKI, Laurent RASSINOUX, Coralie WYSOCZYNSKI, Valérie GADRAT

Actions de communication

- ***Opidor Tour, retour sur une belle aventure,*** article publié sur le site web Inist (lire p. 42).

Les rencontres et interventions sont un levier important pour promouvoir la culture des données FAIR (Faciles à trouver, Accessibles, Interopérables, Réutilisables).

Pour développer les pratiques FAIR : **Partage des données - OPIDoR**

Le portail OPIDoR met à disposition de la communauté de l'enseignement supérieur et de la recherche un ensemble d'outils et de services facilitant la gestion et la valorisation des données afin de répondre aux critères d'intégrité, de reproductibilité et aux principes FAIR qui visent à rendre les données Faciles à trouver, Accessibles, Interopérables, Réutilisables.
Inscrits dans le plan d'actions du plan Données de la recherche du CNRS, les services OPIDoR constituent des composantes essentielles de l'écosystème technique FAIR des données de la recherche.

Le fait majeur

Publication de nouveaux modèles de DMP dans DMP OPIDoR : cc-IN2P3, ENSAM, ICM, PACEA, MASA, Université Toulouse Jean Jaurès, Université Paris Dauphine-PSL, PRODIG, Science Europe, ANR version anglaise.
Démarrage du soutien à la rédaction de DMP : dans le cadre de sa politique science ouverte, l'ANR demande l'élaboration d'un plan de gestion des données (PGD) pour les projets financés à partir de 2019. L'Agence a travaillé en concertation avec l'Inist pour que son modèle de PGD soit intégré dans le portail DMP OPIDoR, pour une saisie en ligne.
Au cours de l'année, l'Inist a accompagné les équipes de recherche du CNRS dans la rédaction de leur plan de gestion de données.

Les autres évènements de l'année...

Poursuite des projets maDMP et maDMP 4LS :
L'objectif de ces deux projets est de rendre les plans de gestion de données (ou Data Management Plans, DMP), produits via l'application DMP OPIDoR, lisibles non seulement par les humains mais aussi par les machines et ainsi automatiser certaines étapes de renseignements des DMP et de mise en œuvre de la gestion des données.

MaDMP 4LS : l'Institut français de bioinformatique (IFB) souhaite par ce biais simplifier le suivi de la gestion et du stockage des données, mais aussi soutenir la production de métadonnées de bonne qualité afin d'accroître son degré de conformité avec les principes FAIR. Au-delà de ce soutien technique, la communauté IFB bénéficiera d'outils de formation pour acquérir de nouvelles compétences en termes de gestion et gouvernance des données.



DOI DataCite : dans le cadre du nouveau modèle économique DataCite, l'Inist a mis en place une nouvelle configuration du consortium français qu'il pilote.

Annuaire CNRS : dans le cadre de l'action 5 de la feuille de route Science ouverte du CNRS, l'Inist participe à la création d'un répertoire des dépôts et des services de données dont le CNRS est responsable et auquel il participe. Ce répertoire permettra d'afficher un élément important de la contribution du CNRS à la science ouverte.

Partenariats et collaborations

- **Institut français de bioinformatique (IFB) et Gesco** : l'Inist est partenaire du projet maDMP 4LS (projet ANR Flash).
- **ANR, GENCI** : partenaires du projet maDMP (machine actionable DMP) de l'unité.
- **ANR, CC IN2P3, ICM - Institut du cerveau et de la mœlle épinière, INRAE, MASA, PRODIG, Science Po** : publication, adaptation de leurs modèles de plan de gestion de données et des recommandations associées dans DMP OPIDoR.
- **MITI, RBDD, GRICAD** : l'Inist participe aux actions de sensibilisation et de formations organisées par les réseaux métiers sur la gestion des données de recherche, le plan de gestion de données.
- **Université de Lorraine** : l'Inist participe au comité opérationnel des données de recherche qui a travaillé en 2020 à la réflexion sur la mise en place d'un entrepôt de données au sein de l'université.
- **DataCite** : membre du consortium, l'Inist est agence d'attribution de l'identifiant pérenne DOI en France.
- **RDA** : l'Institut est membre des working group « exposing data management Plans, Active Data Management Plans IG et FAIR Data Maturity Model: core criteria to assess the implementation level of the FAIR data principles ».
- **UC3-DCC** : l'Inist participe au développement de code commun DMP Roadmap / active DMP.

L'ensemble des services OPIDoR (DMP, CAT OPIDoR et attribution des identifiants DOI) sont signalés dans le plan « données de la recherche » du CNRS inscrit dans la feuille de route pour la science ouverte.

Actions de communication

- ***Le plan de gestion de données et les identifiants DOI : composantes essentielles de l'écosystème FAIR***, Claire FRANÇOIS, Science Ouverte 2020, où en sommes-nous ?, 16 novembre 2020, Paris.
- ***Annuaire des dépôts et des services de données au CNRS : avancement des travaux du groupe de travail***, Paolo LAÏ, Françoise GENOVA (RDA), Science Ouverte 2020, où en sommes-nous ?, 16 novembre 2020, Paris.
- **11 webinaires** de juin à décembre 2020.



DMP OPIDoR : 33 modèles
6000 comptes ouverts
40 établissements de recherche et
jusqu'à plus de 4 000 connexions par mois

Cat OPIDoR : 480 pages créées dans le wiki
368 entités ajoutées (structures ou services)

DataCite : 18 nouveaux adhérents
50 367 DOI créés

Pour tout savoir sur les données FAIR : **Formation - DoRANum**

Le service Formation-DoRANum propose une offre de formation hybride sur la gestion et la conservation, la diffusion et la réutilisation des données scientifiques liées aux activités de recherche.
La plateforme DoRANum, réalisée en partenariat avec le Gis Réseau Urfist, met à disposition près de 100 ressources d'autoformation qui accompagnent les chercheurs à chaque étape du cycle de vie des données. Ce dispositif est complété par une offre de formations à distance (webinaires et classes virtuelles) et en présentiel (ateliers et formations) animées par les équipes.



Le fait majeur

Trois nouvelles ressources pédagogiques ont été créées pour DoRANum :

- « **Parcours interactif sur la gestion des données de la recherche** » : un parcours pédagogique interactif et ludique de 2 h 30 qui a rencontré un vif succès et qui a été repris sur plusieurs plateformes universitaires.
- « **Grille de relecture de plans de gestion de données - Modèle ANR** » : un outil permettant aux équipes de recherche de vérifier la présence des éléments importants dans leur plan de gestion de données.
- « **Tutoriel sur l'outil de rédaction DMP OPIDoR** » : un tutoriel complet de 1 h 30 pour connaître toutes les fonctionnalités de l'outil de rédaction de plan de gestion de données DMP OPIDoR.

Les autres évènements de l'année...

5 ressources externes ont été intégrées dans la plateforme :

- Extrait d'un MOOC sur l'éthique de la recherche produit par l'Université de Lyon ;
- Webinaire des Tuto@mate « À qui appartiennent les données ? » ;
- Trois plaquettes du Réseau des acteurs du Développement LOGiciel ;
- Cours de Teresa Gomez-Diaz « Les logiciels de la recherche et leurs licences : trois visions sur un objet » ;
- « Archivage Numérique des Données de Recherche » interventions captées lors de la journée Archivage Numérique des Données de Recherche organisée le 20 novembre 2019 par l'UMS GRICAD

Mise en ligne de la **première thématique disciplinaire** (physique-chimie) dont les ressources ont été réalisées en collaboration avec la plateforme DATACC (Dispositif d'accompagnement sur les DATA).

Poursuite de **trois projets d'envergure nationale et internationale** :

- **Projet e-learning MIPNES** (Mission de la pédagogie et du numérique pour l'enseignement supérieur) : parcours pédagogique e-learning certifiant sur la gestion, le partage et l'ouverture des données à destination des doctorants, et intégré dans les cursus des écoles doctorales.
- **Projet OBERRED** (Open Badge Ecosystem for the Recognition of skills in Research Data management and sharing) : projet européen Erasmus+ qui a pour objectif de créer un écosystème d'Open Badges au niveau européen pour la reconnaissance des compétences sur le partage et la gestion des données de la recherche dans le cadre de l'Open Science.
- **Plateforme DATACC**, Accompagnement à la gestion des données de recherche en physique et en chimie (Université Lyon 1, Université Grenoble Alpes) : appui pédagogique pour la création des ressources de formation.

Partenariats et collaborations

- **GIS réseau Urfist, Université de Lorraine, Université Côte d'Azur** : partenariat dans le cadre du projet e-learning MIPNES.
- **Université Côte-d'Azur** (France), **University of Applied Sciences Lübeck** (Germany), **Beuth University of Applied Sciences (Germany), University of Göttingen** (Germany), **World University Service of the Mediterranean** (Spain), **Instytut Badań Literackich Polskiej Akademii Nauk** (Pologne), **Institute of Literary Research of the Polish Academy of Sciences** (Pologne), **Università degli studi di Napoli « Federico II »** (Italie) : partenariat dans le cadre du projet OBERRED.
- **EHESS, INRAE, Réseau Urfist** : groupe de travail sur les formations aux DMP, échanges de supports et de bonnes pratiques autour des plans de gestion de données de la recherche contenu multimedia (vidéo, tuto).
- **Huma-Num, GRICAD Grenoble, ESPCI Paris, Université de Lorraine, Université Paris 08, Université Côte d'Azur, Université d'Artois** : partenariats dans le cadre de la mise en place de formations.
- **Université Lyon, Université Grenobles Alpes** : partenariat dans le cadre de la plateforme DATACC.

Webinaires, formations en présentiel et à distance (hors OPIDoR Tour) : près de 30 actions de formation avec + de 1 180 participants

Site web : 44 609 visites
3 759 téléchargements de ressources
79 933 pages vues
76 billets postés
93 ressources pédagogiques

Twitter : 1 032 abonnés / 581 tweets

Youtube : 35 vidéos disponibles
16 000 vues
150 abonnés



L'Inist accompagne la communauté scientifique sur les enjeux de la science ouverte en proposant des dispositifs de formations mixtes.



Le tweet



Pour rendre visibles les données : **Exposition de données**

Le service Exposition de données valorisent les données de recherche sous forme de bibliothèques numériques ouvertes sur le web. Création ou reprise de bases de données anciennes, reformatage et curation de métadonnées en masse, numérisation de documents, les services proposés reposent sur un traitement des données et des métadonnées dans le respect des formats standards et des préconisations du plan national pour la science ouverte. Deux types de service sont disponibles : l'un pour la mise en ligne de grandes bases de données documentaires, l'autre destiné aux laboratoires qui souhaitent mettre en ligne des données de recherche via la création de bibliothèques numériques avec le logiciel OMEKA.

Le fait majeur

Création de la ressource « OpenBibArt » : base de données bibliographique en histoire de l'art en libre accès réalisée en partenariat avec l'INHA (Institut national en histoire de l'art) et le Getty Research Institut (US) en vue de constituer une base de données internationale en histoire de l'art avec un seul point d'accès unifié.

Couvrant la période 1910 à 2007, et avec plus de 1,2 million de références bibliographiques, cet outil va permettre à tout public d'accéder en libre accès à des catalogues numérisés référençant un siècle de recherche en histoire de l'art occidental.

Les autres évènements de l'année…

Bibliothèques numériques OMEKA :

- Création d'un nouveau modèle d'édition numérique sous forme d'un **livre augmenté interactif**, avec la Maison méditerranéenne des sciences de l'Homme (MMSH) et plusieurs laboratoires en sciences humaines et sociales du CNRS. (https://palomed-intg.inist.fr)
- Création d'une **plateforme générique** permettant de coupler données de recherche et publications scientifiques afin d'offrir un accès unifié à la documentation et aux données se référant à des sites archéologiques, avec la MSH Monde et le laboratoire ArScAn (Archéologies et sciences de l'Antiquité. (https://paleoorient-intg.inist.fr)
- Réalisation d'une **plateforme web collaborative** ouverte aux citoyens et bénévoles souhaitant participer et enrichir un projet de recherche, avec le laboratoire ArcHiMèdE (Archéologie et histoire en Méditerranée et en Europe). (https://paleograndest-intg.inist.fr)

Évolution des plateformes Omeka :

- **OMEKA-S,** nouvelle version du logiciel permettant d'intégrer des outils et des approches issues du web sémantique, de connecter plusieurs plateformes entre elles et avec des ressources telles que des thésaurus (PACTOLS) ou des bases de données proposant des formats basés sur les standards du web sémantique.
- **Ajout de services associés à la géolocalisation des données** et à leur éditorialisation sur des fonds de cartes interactifs et variés, en partenariat avec l'IGN.

Partenariats et collaborations

- **Universités de Lyon et de Grenoble-Alpes :** DATACC, projet national retenu dans le cadre de l'appel à projet « Collex-Persée ». Le service Exposition de données met à disposition ses plateformes Omeka dans le but de créer des bases de données de résultats négatifs en physique et en chimie, et de permettre d'accéder aux données, aux protocoles expérimentaux et aux publications associées.

Actions de communication

- ***Omeka, pour des bases de données valorisées***, article publié sur le site web Inist (lire p. 40).
- Poster ***Avec Omeka, une offre de service de l'Inist-CNRS pour des données de recherche valorisées*** présenté lors des Journées Omeka, 13 et 14 octobre 2020, Nancy.



L'atelier de numérisation de l'Inist a traité plusieurs collections de données de recherche, le tout représentant plus de 53 000 pages et documents iconographiques publiés dans les bibliothèques numériques.



En utilisant Omeka-S, le service Exposition de données a étoffé son offre de service pour la science ouverte en enrichissant les données et en les rendant faciles à trouver.

Analyser et fouiller l'information scientifique



Pour des données enrichies et faciles à trouver : **Text and data mining**

Le service *Text and data mining* a pour objectif de développer des méthodes de fouilles de textes afin d'enrichir les données présentes dans des réservoirs ouverts comme Conditor ou ISTEX, mais également de mettre à disposition ces méthodes à des utilisateurs non spécialistes. La constitution de la base Conditor fait aussi partie du service.



Le fait majeur

Ouverture officielle en février de la plateforme Conditor lors d'une journée de présentation au MESRI. Conditor vise à recenser la production scientifique des unités et établissements de l'ESR, à consolider et à enrichir les données dédoublonnées afin de les mettre a disposition pour, entre autres, des analyses bibliométriques en mettant l'accent sur les publications en accès libre.
Conditor est compatible avec l'application LODEX permettant d'obtenir des tableaux de bord visuels.

Actions de communication

- ***Enrichissement automatique de données bibliographiques : retours d'expériences***. Pascal CUXAC, Wikicite 2020 « Wikicite et scientométrie ouverte », 27 octobre 2020.
- ***La fouille de textes aujourd'hui... et demain ? résultats de l'enquête.*** Fabienne KETTANI, Billet de Blog.
- ***Atelier TextMine – Fouille de textes***, Pascal CUXAC et Vincent LEMAIRE, EGC 2020, Bruxelles, janvier 2020.
- ***Conditor aujourd'hui***, Valérie BONVALLOT, Conditor, un service de référencement opérationnel en 2020, 6 février 2020, MESRI Paris.
- ***ISTEX et Conditor convertis au format TEI***, article publié sur le site web Inist (lire p. 39).

Les autres évènements de l'année...

Conception d'un **algorithme de classification** apposant des domaines scientifiques sur les signalements bibliographiques des documents.

Mise au point de **méthodes d'homogénéisation d'affiliations** à partir du Répertoire national des structures de recherche (RNSR).

Mise à disposition d'un **moissonneur** permettant d'extraire des corpus de Conditor et de générer dynamiquement une notice « de référence » unique quand on détecte des doublons.

Détection des articles en accès ouvert via un alignement avec les bases CORE et Unpaywall, ainsi qu'un alignement avec le Baromètre de la science ouverte (BSO).

Partenariats et collaborations

- **Université de Toulon** : collaboration pour permettre la constitution de tableaux de bord à partir de Conditor.
- **AMUE (Agence de mutualisation des universités et établissements)** : intégration de Conditor dans le projet Caplab de l'AMUE afin d'associer les publications aux projets de recherche.
- **Université de Limoges** : le système de pilotage de la recherche de l'Université de Limoges, dans le cadre de son projet YUZU, alimente et enrichit ses rapports d'analyses décisionnelles par des données Conditor.

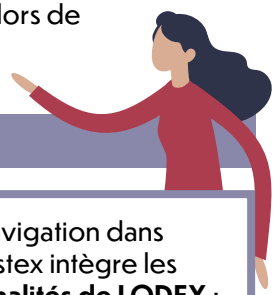
Conditor :

1 430 256 notices
1 108 297 notices uniques
4 sources

Alliant compétences d'infrastructures techniques, terminologiques et de connaissances scientifiques, l'Inist développe des outils et services d'exploration des contenus, en adéquation avec les demandes des chercheurs.

Pour faciliter la fouille de textes : **Textes & corpus - ISTEX**

L'équipe ISTEX poursuit ses actions de développement de la plateforme ISTEX et des services associés pour faciliter leur utilisation dans des travaux de fouille de textes : amélioration des performances et de l'ergonomie des services ; cycle de formations en France pour la constitution de corpus ; communication autour de la constitution de corpus spécialisés lors de conférences telles que LREC et TALN...



Le fait majeur

ISTEX-DL, le service de téléchargement associé à ISTEX, a fait peau neuve et a repoussé ses limites :

- augmentation des possibilités de téléchargement (jusqu'à 100 000 documents) ;
- des requêtes sans limite de taille ;
- visualisation d'un échantillon de résultats ;
- parcours de téléchargement par type d'usage ;
- une connexion facilitée avec l'outil LODEX ;
- un nouveau format de téléchargement (tar.gz) ;
- une interface adaptée aux mobiles (responsive).

Les autres évènements de l'année...

Data.istex propose de **nouvelles vues** sur les données ISTEX :

- la **revue de sommaire** permet de télécharger en format Kbart la couverture réelle d'ISTEX afin d'avoir une vue globale ou partielle de l'état de collection des livres et des journaux électroniques ;
- la **nouvelle rubrique « Corpus d'actualité »** invite à une exploration du passé pour éclairer les problématiques du présent (ex. : SARS COVID) ;
- le **jeu de données « Ayants-droit à l'usage d'ISTEX »** propose un alignement des données sur les ayants droit de l'ESR avec les données d'autorité de IdRef ;
- de **nouveaux corpus spécialisés** sont mis à la disposition de la communauté scientifique spécialisée en fouille de textes :
  - **Animalia 100** : un corpus de 100 documents en zoologie enrichis avec des noms d'espèces animales détectés automatiquement à l'aide de trois outils (entity Fishing & IRC3, Ty+Rex) ;
  - **Beethoven** : un corpus de 1800 documents en musicologie qui a pour objectif d'illustrer la méthodologie de constitution d'un corpus pour une application en fouille de textes.

Afin de faciliter la navigation dans ses données, data.istex intègre les **nouvelles fonctionnalités de LODEX** :

- un **fil d'Ariane** pour se repérer entre les différentes pages ;
- une **barre de navigation** pour accéder aux différentes fonctionnalités, parmi lesquelles le bouton « Télécharger » en lien direct avec ISTEX-DL ;
- une **interface responsive** adaptée aussi bien sur mobile que sur ordinateur ;
- de **nouvelles représentations graphiques** pour retrouver des données similaires au fil de sa navigation.

Grâce à data.istex, les chercheurs peuvent accéder à des corpus librement et simplement.

Partenariats et collaborations

- **Abes :** négociation de nouveaux corpus afin d'enrichir le fonds ISTEX en vue d'augmenter sa couverture temporelle par de nouvelles années pour le corpus EDP Sciences (2017, 2018, 2019) ainsi que sa couverture scientifique, par l'intégration de nouveaux corpus éditeurs en Sciences humaines et sociales (Droz, Open Edition (ebooks), Taylor & Francis), en biologie et histoire des sciences (RSL).

30 corpus éditeurs chargés
12 corpus spécialisés pour la fouille de textes
1 corpus d'actualité (Coronavirus)
9 314 revues / **348 636** monographies
23 351 350 documents dont
22 903 845 documents avec au moins **1** enrichissement
83 213 426 enrichissements générés sur les documents



Actions de communication

- ***An Experiment in Annotating Animal Species Names from ISTEX Resources***. Sabine BARREAUX, Dominique BESAGNI, 12th Conference on Language Resources and Evaluation (LREC 2020), May 2020, Marseille, France. pp.4509 - 4513. hal-03008209.

- ***Vers un corpus optimal pour la fouille de textes : stratégie de constitution de corpus spécialisés à partir d'ISTEX (démonstration)***. Camille DE SALABERT, Sabine BARREAUX. 6e conférence conjointe Journées d'études sur la parole (JEP, 33e édition), Traitement automatique des langues naturelles (TALN, 27e édition), Rencontre des étudiants chercheurs en informatique pour le traitement automatique des langues (RÉCITAL, 22e édition). Volume 4 : Démonstrations et résumés d'articles internationaux, Juin 2020, Nancy, France. pp.66-69. hal-02768520v3.

- **Ateliers pratiques autour d'ISTEX** : poursuite du tour de France en collaboration avec le réseau des URFIST pour accompagner les utilisateurs dans la constitution d'un corpus spécialisé à partir du réservoir ISTEX et dans sa valorisation à l'aide d'outils pour la visualisation et la fouille de textes (LODEX, Gargantext, Cillex) **4 sessions à Paris, Montpellier, Strasbourg** ; animateurs : Claude NIEDERLENDER, Camille DE SALABERT, Pascale VIOT, Sabine BARREAUX ; 38 stagiaires parmi lesquels des professionnels de l'IST, des doctorants, des enseignants-chercheurs.

- **14 billets** publiés sur le blog ISTEX.

- ***Lodex : quand la science rencontre la culture***, article publié sur le site web Inist (lire p. 41).

- ***ISTEX-DL pour un accès facilité aux publications scientifiques***, article publié sur le site web Inist (lire p. 43).



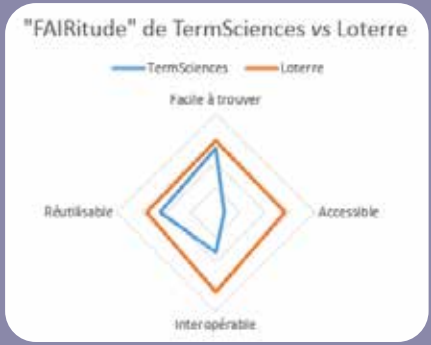
Pour un traitement optimal des terminologies : **Ingénierie terminologique**

Les activités d'ingénierie terminologique se sont poursuivies en 2020, alliant le conseil, la conception et la mise en ligne de nouvelles ressources terminologiques sur Loterre, la plateforme terminologique en libre accès de l'Inist.

Le fait majeur

En réponse à une proposition du projet FooSIN, le service a évalué le **positionnement de la plateforme Loterre** vis-à-vis du respect des principes FAIR en utilisant la grille SHARC développée par le groupe RDA SHARC.
Cette grille de notation comprend une série de questions portant sur les méthodes de représentation des données, les techniques de stockage ou les moyens mis en place pour le partage.

Cette métrique a été appliquée non seulement à Loterre mais aussi à son prédécesseur, **TermSciences**. Cette évaluation comparée a permis de prendre la mesure de l'évolution des pratiques d'exposition de ressources sémantiques de l'Inist au cours des 15 années qui séparent la mise en œuvre des deux plateformes, pratiques aujourd'hui résolument orientées vers la science ouverte et la prise en compte des principes de partage et de réutilisation des données.



Les autres évènements de l'année...

En réponse aux appels de la communauté scientifique visant à partager les connaissances liées à la pandémie de **COVID-19**, un thésaurus bilingue (français-anglais) a été créé et exposé sur Loterre.

Lancement du projet « **Paléosaurus** », en collaboration avec des chercheurs et ingénieurs de l'Université Paris-Saclay et du LSCE (Laboratoire des sciences du climat et de l'environnement). L'objectif consiste à créer un vocabulaire normé en paléoclimatologie, qui fait défaut au plan international et qui sera en mesure de devenir une référence pour l'ensemble de la communauté scientifique et une source de métadonnées pour l'indexation des données de la recherche du domaine.

L'ouvrage ***Modélisation ontologique & Psychologies : une influence réciproque*** (Éditions Matériologiques), co-rédigé par Frank Arnould (Inist) et Xavier Aimé (Cogsonomy) a été finalisé à l'automne 2020.
Il est disponible sur le site internet de l'éditeur.

Complètement ouverte et FAIR, la plateforme Loterre est résolument tournée vers la science ouverte.

Partenariats et collaborations

- **Laboratoire des sciences du climat et de l'environnement (LSCE) et Learning Center de l'Université Paris-Saclay :** projet de création d'un thésaurus de paléoclimatologie.
- **Muséum national d'Histoire naturelle (MNHN) :** échanges visant à la publication sur Loterre d'un thésaurus relatif à la nomenclature zoologique.
- **BnF – Équipe Rameau :** travaux d'enrichissements réciproques des terminologies relatives à la taxonomie des poissons.
- **Groupement de recherche Mémoire (GDR 2013) et SémanDiv (GDR 2011) :** appui en ingénierie terminologique.

Ajout de **7** ressources terminologiques :

- **3** ressources de l'INRAE : Lexique de pédologie, Technologie des semences, Agriculture et systèmes d'élevage
- **4** ressources produites par l'Inist : Ethnologie, Pathologies humaines, COVID-SRAS, Electrotechnique-Electroénergétique

Loterre : **40** terminologies (**+ 21 %**)
454 832 concepts (**+ 6,4 %**)
5 258 408 triplets (**+ 7,2 %**)



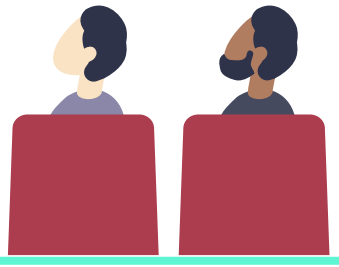


Le tweet



Actions de communication

- ***Évaluation « FAIRitude » pour Loterre et TermSciences***, article publié sur le site web Inist (lire p. 43).



Pour un suivi sur les publications : **Appui au pilotage scientifique**

L'équipe Appui au pilotage scientifique réalise des études bibliométriques permettant de caractériser l'activité de publication d'un laboratoire, d'un institut ou de toute autre structure, ou celle liée à une thématique scientifique : évolution du nombre de publications, mise en évidence des collaborations, principales thématiques, revues les plus utilisées, etc.

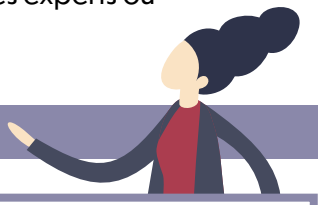
Parallèlement, des prestations de veille documentaire sont proposées pour dresser un panorama thématique, repérer des experts ou encore pour rechercher des partenaires sur un sujet précis.

Le fait majeur

Le service a quantifié les **publications scientifiques en accès libre produites par le CNRS**. 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 ».

Le service Appui au pilotage scientifique a vocation à fournir chaque année les chiffres du baromètre de la science ouverte (BSO) du CNRS.

Les autres évènements de l'année...

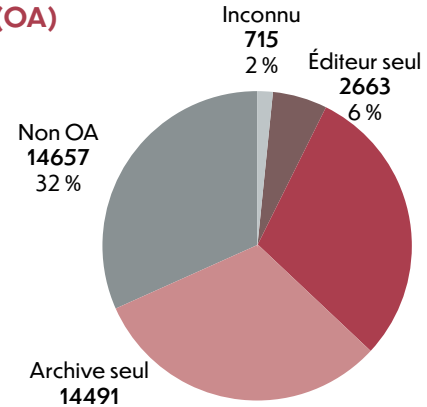


Démarrage de la réalisation des études concernant les **six grands défis sociétaux** identifiés par le CNRS dans le Contrat d'objectifs et de performance 2019-2023. La première étude concerne le **changement climatique.**

Participation au développement de **LODEX**, outil libre permettant de mettre en ligne des données structurées, notamment via des représentations graphiques et offrant un accompagnement des utilisateurs de l'ESR, du Cirad, de l'Inrae et du BRGM.

Étude des publications CNRS en Open Access (OA)

Réalisé avec l'outil **LODEX**, ce graphique présente la part des publications disponibles en accès libre sur 46 308 publications produites en 2019 par le CNRS.
Il propose la répartition selon le moyen d'accès : via le site de l'éditeur, à travers une archive ouverte, ou les deux (commun).
(données WoS)





Partenariats et collaborations

- **SAPPS** : poursuite du partenariat pour produire les indicateurs bibliométriques du CNRS (hébergement des données utilisées, repérage des publications du CNRS, réflexion sur de nouveaux indicateurs).
- **Université de Lorraine (UL)** : démarrage d'une collaboration avec la mission inter-services Bibliométrie de l'UL récemment créée, pour un échange des pratiques, une présentation et la mise à disposition de l'outil LODEX.

ILS TÉMOIGNENT...

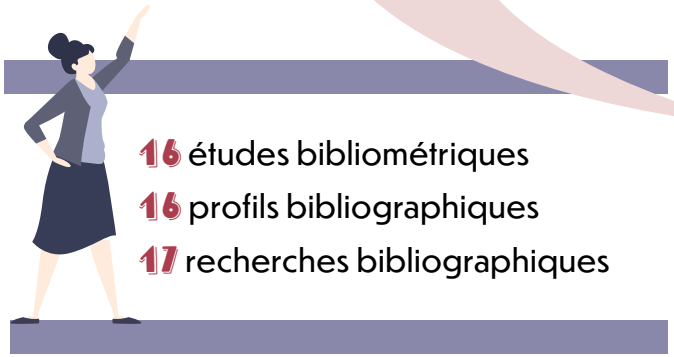
Merci beaucoup pour ces indicateurs. Cela me permet de préparer des planches en vue de la visite de la Ministre des outremers la semaine prochaine...
Annaïg LE GUEN CNRS
Vice-Présidente du RESIPOL
Directrice de l'USR mixte CRIOBE (CNRS, EPHE, UPVD)
Centre de recherches insulaires et observatoire de l'environnement - LabEx CORAIL

Nous avons remis au CNRS fin janvier un document intitulé « Bilan et Défis des Spectroscopies d'électrons » à l'Institut de chimie du CNRS, qui a intégré des données bibliographiques établies grâce à votre aide. Ce document a été examiné par l'INC et a été très apprécié.
Hervé MARTINEZ
Professeur des Universités
Directeur Adjoint Recherche Innovation Collège STEE
Sciences et technologies pour l'énergie et l'environnement
IPREM UMR5254 - UPPA/CNRS

Focus 16 études bibliométriques

Quelques exemples :

- Mise en évidence, pour la période 2016-2018, de la **production des laboratoires de l'Institut de physique** en lien avec 16 TGIR français, étrangers ou internationaux dans les domaines des sources de neutrons et de rayonnement synchrotron, des lasers et champs intenses.
- Suivi de la **production française dans le domaine de l'énergie** entre 2014 et 2019 à la demande de l'Institut des sciences de l'ingénierie et des systèmes.
- Mise à jour du suivi de la **production du TGIR ESO European Southern Observatory** pour la période 2018-2020 à la demande de l'Institut national des sciences de l'Univers.
- **Évaluation du nombre de publications** des laboratoires d'outre-mer pour la période 2010-2020 à la demande de l'Institut écologie et environnement.
- **Suivi de la production** de différents laboratoires ou structures du CNRS : Institut des sciences biologiques, Institut des sciences de l'ingénierie et des systèmes, OSU OTELo, CRIOBE – USR3218, LEM UMR – 9221, PACEA – UMR5199, etc.



Actions de communication

- ***La production du CNRS en libre accès quantifiée par l'appui au pilotage scientifique***, article publié sur le site web Inist (lire p. 39).

Pour un appui linguistique : **Traduction**

Le service propose des traductions et relectures de communications scientifiques des laboratoires et institutions.

Le fait majeur

Traduction du **rapport du groupe Knowledge Exchange** intitulé « *Open Scholarship and the need for collective action* » paru en octobre 2019 dans la version anglaise. Ce rapport de 100 pages comporte près de 39 000 mots. Il explore l'économie de la science ouverte et propose une nouvelle approche dans six pays européens (Danemark, Finlande, France, Allemagne, Pays-Bas et Royaume-Uni) basée sur le rôle crucial des communautés de recherche.

Les autres évènements de l'année...



Traduction de « **Une stratégie mondiale pour la science ouverte. Proposition de la France sur la science ouverte dans le cadre de la consultation de l'Unesco** », contribution du ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation dans le projet de recommandation sur la science ouverte par l'Unesco.

Le recours à la **traduction automatique neuronale post éditée** confirme l'incursion massive de l'IA (intelligence artificielle) dans l'exercice du métier et son appropriation dans l'ensemble de l'équipe.

Parmi les **autres travaux de traduction** : consignes du site ESGBU (Enquête statistique générale des bibliothèques universitaires) ; site internet de l'Institut Jean Lamour à Nancy, contrats de licence éditeur ; cinq articles destinés à être publiés dans la revue SET (Sciences, eaux & territoires), à la demande de l'Inrae.

Contenus multimédia

Relecture, pour le service Formation - DoRANum de l'Inist, de 12 chapitres du **MOOC « Foster Open Science »**, soit 144 pages illustrées (près de 42 000 mots).

Traduction en anglais d'une centaine de **descriptions de films pour la vidéothèque du CNRS** sur des sujets scientifiques variés (ethnologie, histoire des sciences et techniques, histoire du cinéma scientifique...).

106 demandes traitées
732 748 mots



En plus des missions ponctuelles de traduction pour des documents liés à la science ouverte, les traducteurs interviennent tout au long de l'année pour le site *Ouvrir la science* du Comité pour la science ouverte.

L'Inist, c'est aussi...



Les activités de l'Institut s'articulent autour de trois axes forts structurés par des équipes opérationnelles.
Ces dernières sont épaulées au quotidien dans leurs missions par des équipes dites fonctions transverses associées **au secrétariat général, aux systèmes d'information, à la communication, au web et à la veille**.

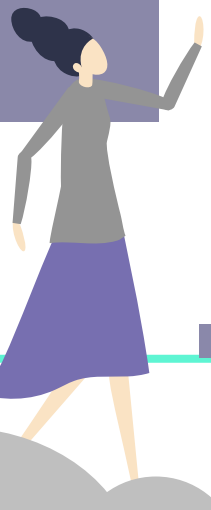


Pour le fonctionnement de l'Institut : **Secrétariat général**

Le fonctionnement de l'Institut s'appuie sur une gestion rigoureuse et prospective nécessaire au suivi de l'activité. Le secrétariat général englobe les volets RH (ressources humaines), financier, réglementaire, logistique et technique.

Le fait majeur

Alors qu'il assurait certaines missions bâtimentaires relevant du « propriétaire »,l'Inist est devenu « **locataire** » à part entière au 1er janvier 2020.
Ce changement organisationnel et fonctionnel a nécessité un important travail relatif au transfert des activités administratives, liées à la maintenance du bâtiment, vers la Délégation régionale Centre-Est du CNRS (clôture et solde de contrats).



Les autres évènements de l'année...

Afin d'améliorer les **conditions de travail** du personnel les stores ont été changés, un rideau d'air chaud a été posé et certaines toilettes ont été réfectionnées. De plus, les salles de réunions ont été équipées de grands écrans.

L'Institut a accueilli **trois nouvels agents** recrutés en concours externes.

Effectif global : 154 agents et 9 stagiaires
Budget : 5 695 000 euros
dont 4 498 000 euros pour les acquisitions d'abonnements électroniques
3000 heures de formation pour environ 100 agents
1500 bons de commande
50 contrats
10 marchés à procédure adaptée (MAPA) menés

Pour répondre aux besoins de sécurité : **Système d'information**

Le Système d'information (SI) garantit la continuité du service informatique fourni aux utilisateurs. Moderne et performant, il répond notamment aux besoins de sécurité.

Le fait majeur

L'arrivée soudaine de la **crise sanitaire** a nécessité la mise en œuvre rapide d'outils (système de visioconférence, tchat, généralisation du VPN) et de matériels (80 PC portables) pour que l'activité en télétravail puisse être efficace.

Les autres évènements de l'année...

Renforcement de la sécurité du SI :
- changement du firewall,
- réécriture des règles de filtrage,
- répartition optimisée des équipements dans notre datacenter.

Installation d'une **nouvelle baie de stockage** afin de mutualiser les solutions de stockage « Inist » et « ISTEX » devenues obsolètes.

Migrations de logiciels vers les dernières versions (VMWare, Veeam, Dataprotector).



L'optimisation du SI pour encaisser près de quatre millions de requêtes par jour est une démonstration de son ouverture pour l'accès à la science.

600 serveurs virtuels et **35** serveurs physiques
850 To de capacité brute de stockage
380 bases de données
300 postes de travail
3 robots de sauvegarde (capacité de **3** Po)
185 environnements docker
4 millions de requêtes par jour
99,8 % d'applications disponibles

Pour promouvoir les projets et services : **Communication - Web**

Le service Communication-Web a pour mission de valoriser et de promouvoir les services et projets de l'Inist. À travers une stratégie de communication, il met en place des actions de communication en étroite collaboration avec les équipes de l'institut.
Il conçoit aussi des sites web et des supports de communication.



Le fait majeur

La **crise sanitaire** a eu un impact important sur les actions de communication en externe, plusieurs rencontres ayant été annulées.
La communication interne a quant à elle eu un rôle important à jouer.
Une charte graphique inspirée du style sketchnote a été réalisée en interne, en lien avec les chargés de prévention, pour l'affichage des règles sanitaires en vigueur.
La généralisation du télétravail a aussi induit la mise en œuvre de nouveaux moyens de communication principalement autour de visio-conférences.

Après avoir participé à la conception du site *Ouvrir la science* du Comité pour la science ouverte hébergé à l'Inist, le service Web gère sa maintenance.

Les autres évènements de l'année...

Six vidéos de présentation ont été réalisées en lien avec le service Formation de l'Institut : les trois axes de l'Institut, BibCnrs, le bilan Opidor Tour, ISTEX-DL.
Elles sont visibles sur la chaîne Youtube « Inist-CNRS ».

Comme chaque année, l'Inist a collaboré avec la Délégation régionale Centre-Est du CNRS à l'organisation de l'événement multipartenarial « **Les Sciences sur la Place** », dans le cadre du Livre sur la Place à Nancy.

Twitter : 3 974 abonnés (+9,6 %)
193 tweets
466 800 impressions
8 558 visites du profil

Site Inist : 70 789 visites
435 255 pages vues
10 articles « Réalisations » publiés
72 actualités publiées

Web : 41 sites en productions dont 4 nouveaux



Actions de communication

- Présence de l'Inist dans l'espace numérique dédié aux présentations des directions fonctionnelles au côté de la DIST (Direction de l'information scientifique et technique), **Journée nationale d'accueil des nouveaux entrants du CNRS**, 17 septembre 2020, Paris.

Pour le suivi de l'actualité : **Veille**

Le service Veille accompagne le monde de la recherche et les professionnels du domaine en mettant à disposition en continu une mine d'informations actualisées sur le domaine de l'information scientifique et technique via les sites LaLIST et Ouvrir la science.

LaLIST
Le site propose chaque jour des actualités en IST à destination des professionnels du domaine, des chercheurs, des enseignants, des doctorants, voire du grand public.
La veille diffusée sur le site LaLIST s'appuie sur le projet Inist dans lequel s'inscrivent les trois axes forts de l'Institut. Parmi les thématiques phares, on retrouve les données de la recherche, le TDM (text and data mining) et tous les autres enjeux de la science ouverte dans le monde entier.

Ouvrir la science
Piloté par le Comité pour la science ouverte (CoSO), le site est entièrement dédié à la science ouverte. Il est alimenté au fil de l'eau par un travail de veille opéré à l'Inist.
Ses objectifs sont notamment de valoriser le travail du CoSO, de fournir des informations, des documents et des outils, et d'exprimer les orientations politiques de la France en matière de science ouverte.

Le fait majeur

Lancement du projet de refonte du site LaLIST avec l'objectif de lui donner un nouveau dynamisme, tout en prenant en compte les préoccupations et centres d'intérêt de la communauté de l'enseignement supérieur et de la recherche.



LaLIST : 3 292 abonnés (tous canaux + 8,75 %)
1 073 billets publiés
57 751 pages vues

Ouvrir la science : 74166 visites (+ 65 %)
149 001 pages vues
15 829 téléchargements
59 articles publiés sur le blog
45 ressources
43 références agenda
9 newsletters

L'Inist développe, maintient et fait évoluer le site Ouvrir la science. Le service gère également tous les contenus de veille et l'agenda.

Partenariats et collaborations

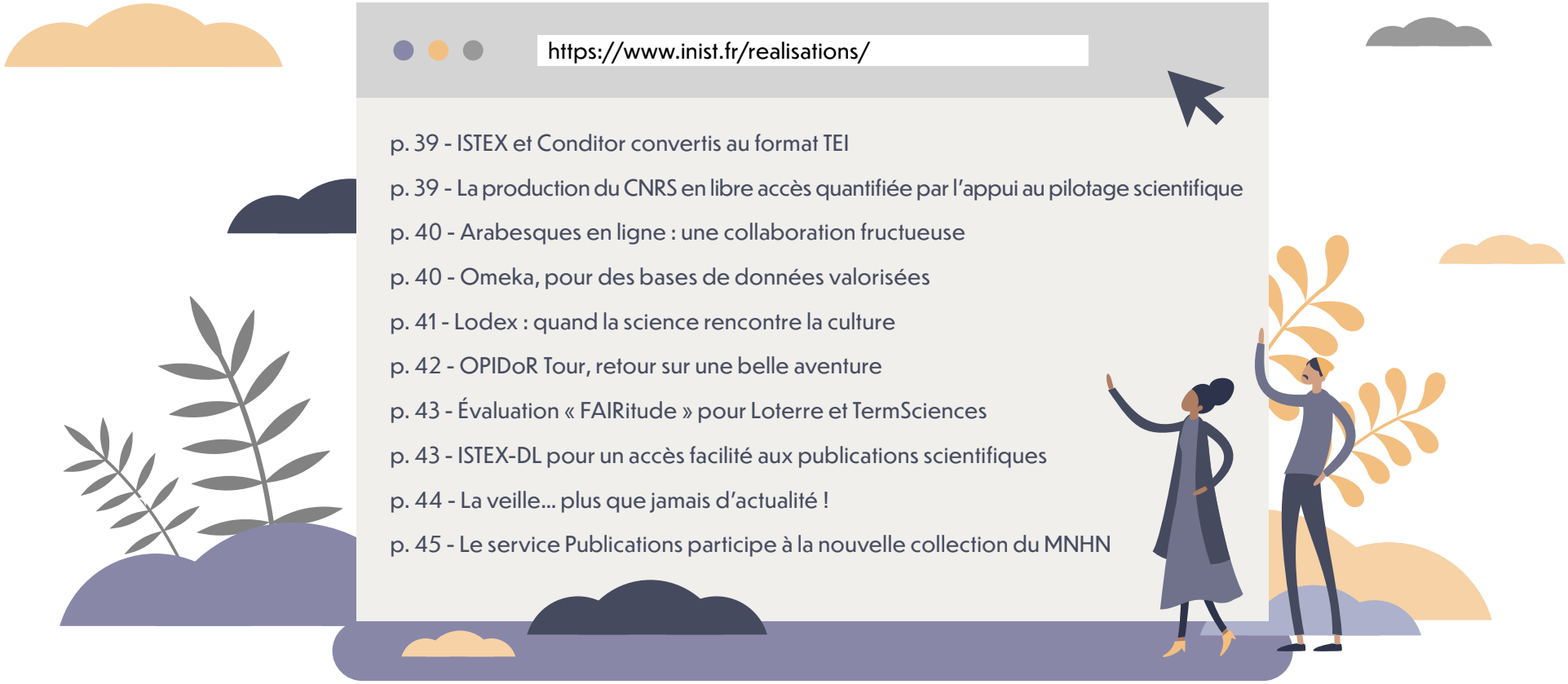
- **CoSO :** collaboration pour le site Ouvrir la science.

Actions de communication

- ***La veille... plus que jamais d'actualité !***, article publié sur le site web Inist (lire p. 44).

Pour en savoir plus...

Articles publiés en 2020 sur le site institutionnel



p. 39 - ISTEX et Conditor convertis au format TEI

p. 39 - La production du CNRS en libre accès quantifiée par l'appui au pilotage scientifique

p. 40 - Arabesques en ligne : une collaboration fructueuse

p. 40 - Omeka, pour des bases de données valorisées

p. 41 - Lodex : quand la science rencontre la culture

p. 42 - OPIDoR Tour, retour sur une belle aventure

p. 43 - Évaluation « FAIRitude » pour Loterre et TermSciences

p. 43 - ISTEX-DL pour un accès facilité aux publications scientifiques

p. 44 - La veille... plus que jamais d'actualité !

p. 45 - Le service Publications participe à la nouvelle collection du MNHN



ISTEX et Conditor convertis au format TEI

Publié le 13 janvier 2020

Afin que les données de chacun des réservoirs soient interopérables avec d'autres systèmes, les équipes ISTEX et Conditor de l'Inist convertissent les données qu'elles récupèrent au format TEI (Text Encoding Initiative).

Tandis qu'ISTEX est un réservoir de ressources multidisciplinaires (texte intégral et notices bibliographiques), Conditor est quant à lui un projet de référentiel national de la production scientifique française collectant des notices bibliographiques issues de plusieurs sources. Ils ont cependant en commun d'être tous deux orientés vers la **science ouverte** et les **principes FAIR** (Facile à trouver, Accessible, Interopérable, Réutilisable).

C'est donc dans ces objectifs d'ouverture et d'interopérabilité que les données ISTEX, initialement au format XML source (format éditeur), et les données Conditor, multiformats, ont été converties en TEI, un format XML de données structurées.

Un format pivot

Les données ISTEX et Conditor se retrouvent respectivement sous un **format pivot TEI** qui leur est propre ; format qui permet d'optimiser la réutilisation des données et qui permet aussi à ces dernières d'être interrogées par les moissonneurs TEI.

Pour ISTEX, 26 corpus de textes intégraux fournis par plusieurs éditeurs (soit 26 formats différents) sont donc passés dans la moulinette de feuilles de style complexes pour en ressortir avec des données classées et rangées. Ces enrichissements ont pour but de permettre aux utilisateurs d'analyser automatiquement des masses importantes de données par le biais de services à valeur ajoutée (enrichissement des données, analyse sémantique, mutualisation des connaissances…).

Pour Conditor, quatre corpus sont actuellement reformatés en TEI. Deux le sont grâce à des feuilles de style, les deux autres à l'aide de programmes développés à l'Inist.

Un schéma ODD-ISTEX

Ces corpus, confirmés par un validateur, répondent désormais au **schéma ODD** (One Document Does it all) qui a plusieurs avantages notamment de fournir un format XML bien établi, de permettre un traitement fortement intégré avec d'autres systèmes TEI, d'offrir une pérennisation à long terme et une standardisation.

Cependant, la TEI n'a pas d'éléments spécifiques pour encoder des informations supplémentaires dites d'enrichissement. Les équipes ISTEX ont donc produit un schéma ODD-ISTEX permettant de documenter et de valider les enrichissements spécifiques produits par les équipes ISTEX et leurs partenaires (ex. : l'enrichissement des entités nommées comme la désambiguïsation et l'alignement).

Plus d'informations sur le schéma ODD-ISTEX (https://blog.istex.fr/schema-odd-istex/)

La production du CNRS en libre accès quantifiée par l'appui au pilotage scientifique

Publié le 3 février 2020

À la demande de la DIST (Direction de l'information scientifique et technique), l'Institut a quantifié les publications scientifiques en accès libre produites par le CNRS en 2017. Avec un résultat de 49 %, le CNRS s'affiche comme un très bon élève face à une production nationale en libre accès qui s'élève à 41 %.

Le travail a commencé il y a quelques mois lorsque Sylvie Rousset, directrice de la DIST, a sollicité l'Inist pour quantifier l'open access dans la production scientifique du CNRS. Pour sa mise en œuvre, l'équipe Appui au pilotage scientifique a collaboré avec l'équipe produisant le baromètre de la science ouverte (BSO) qui a pour objectif de mesurer les progrès de l'accès ouvert aux ressources scientifiques au niveau national et qui s'inscrit dans le cadre du Plan national pour la science ouverte.

Une méthodologie bien précise

La première étape a consisté à repérer la production française 2017 dans la base de données WoS (Web Of Science) qui fait référence en matière de publications scientifiques : le résultat affichait **112 654 notices différentes** (identifiants WoS différents).

Ce corpus a été transmis au ministère de l'Enseignement supérieur et de la Recherche, qui pilote le baromètre de la science ouverte, avec l'information de DOI (présent à 84,5 %) complétée par le titre, la source et l'auteur. Avec l'outil Unpaywall qui permet de retrouver les articles scientifiques en accès libre sur le web, et grâce au DOI, le fichier a ensuite été retourné « marqué » de l'information « Open Access » (oui ou non) pour 99 737 lignes soit 88,5 %.

Afin de mesurer le poids de l'open access dans la production CNRS, et pour le comparer à la production nationale, le périmètre français a été restreint selon les critères utilisés par le **Service d'appui à la politique et à la prospective scientifiques (SAPPS) du CNRS** avec lequel l'Inist travaille au repérage des publications CNRS dans le WoS. Ce périmètre prend en compte deux bases du WoS Core Collection (Science Citation Index Expanded et Conference Proceedings Citation Index) ainsi que certaines typologies de documents.

Un résultat qui dépasse la production scientifique nationale

Après quelques mois de travail opéré par l'équipe de l'Inist, les résultats sont tombés mettant le CNRS au rang des très bons élèves. En effet, **49 % de la production 2017 du CNRS sont disponibles en accès libre**, répartie presque à parts égales entre les sites des éditeurs et les

archives. La production française globale pour 2017 disponible en accès libre est quant à elle de 41 %.

Les résultats produits par l'équipe Appui au pilotage scientifique ont notamment été repris dans le Journal du CNRS n° 298 (décembre 2019) qui a consacré un article sur la science ouverte intitulé « Science ouverte : la révolution nécessaire ».

Et la suite ?

Produire des chiffres, c'est bien, mais les mettre à jour régulièrement, c'est mieux. Aussi, l'Inist prévoit de mettre à jour les données pour 2017 compte tenu du fait que certaines publications peuvent devenir en libre accès ultérieurement.

Il est également prévu de faire ce travail sur l'année 2018. Cette action permettra d'observer l'évolution du nombre de publications scientifiques en accès libre produites par le CNRS et ainsi de constater si la politique de libre accès impulsée par le CNRS dans le cadre de la science ouverte porte ses fruits.

Arabesques en ligne : une collaboration fructueuse

Publié le 7 février 2020

La revue *Arabesques* de l'Abes est désormais en ligne sur Prairial, portail des revues de l'Université Lyon 3. Pour la mise en ligne des anciens numéros, l'Abes a sollicité le service Publications de l'Inist qui s'appuie sur la chaîne éditoriale Métopes - Méthodes et outils pour l'édition structurée.

Avec l'objectif d'améliorer la diffusion en ligne de la revue ***Arabesques***, l'Abes a fait appel à l'Inist afin de rendre disponible en ligne la **collection complète** de la revue sur le portail de revues scientifiques Prairial. Conçu par les BU de l'Université Lyon 3, le portail a obtenu en 2018 le statut de **pépinière officielle d'OpenEdition** et s'inscrit dans le paysage de la **science ouverte**.

Formation à l'utilisation de Métopes

Pour répondre à la demande de l'Abes, le service Publications s'est tout d'abord formé à la chaîne éditoriale Métopes, développée par le Pôle document numérique de la Maison de la recherche en sciences humaines (MRSH) de Caen dans le cadre de **l'infrastructure de recherche Métopes**.

Cette chaîne éditoriale a la particularité de gérer le format XML pour la diffusion multi-supports, assurant ainsi une pérennisation des contenus et une haute qualité des métadonnées associées.

Numérisation, stylage et mise en ligne

Pour ce projet qui concerne les numéros 1 à 91 de la revue, l'équipe travaille sur les PDF éditeurs ainsi que sur les PDF issus de la numérisation. La numérisation des numéros 1 à 17 a été réalisée par le service dédié de l'Inist durant l'été 2019.

À partir de ces fichiers, quatre membres du service Publications réalisent un travail de **stylage des métadonnées** en définissant entre autres les titres, auteurs, affiliations, mots-clés ou encore les tableaux et images. De plus, les données d'autorité sont alignées avec celles de l'application IdRef – Identifiants et Référentiels pour l'enseignement supérieur et la recherche.

Une fois le travail terminé pour chaque article, le dossier contenant le fichier XML enrichi, les images ainsi que le PDF initial, est chargé dans le logiciel d'OpenEdition Lodel utilisé par le portail des revues de l'Université Lyon 3.

À ce jour, 25 numéros d'*Arabesques* (du n° 64 au n° 96) sont disponibles via ce portail. L'ensemble de la revue devrait être en ligne d'ici un an.

Une collaboration fructueuse à double sens

Au-delà d'aider l'Abes à concrétiser ce projet, ce fut, pour le service Publications de l'Inist, l'occasion de développer ses compétences dans l'utilisation de la chaine éditoriale Métopes. Ce solide savoir-faire laisse ainsi entrevoir une future prise en charge des revues d'OpenEdition ou du MNHN avec qui l'Inist est partenaire.

TÉMOIGNAGE

« L'Abes remercie chaleureusement l'ensemble de ses partenaires pour leur implication dans ce projet collaboratif réalisé au profit de la communauté des bibliothèques de l'ESR. Grâce à l'Inist-CNRS, qui a assuré avec efficacité le traitement rétrospectif des anciens numéros, la collection entière sera progressivement mise en ligne. »

Marie-Pierre Roux, Abes

Omeka, pour des bases de données valorisées

Publié le 18 février 2020

Exposer, partager et valoriser vos données de recherche sous forme de bibliothèques numériques, tel est le service proposé par l'équipe de l'Inist chargée de l'exposition des données qui a déjà mis en ligne plusieurs bases de données éditorialisées.

Utilisé pour la **création de bibliothèques numériques, Omeka** est un logiciel libre qui a rapidement séduit les acteurs de la recherche, notamment la communauté Sciences humaines et sociales (SHS). Aujourd'hui, Omeka est un outil qui s'inscrit pleinement dans la **science ouverte** et qui permet la création de bases de données au plus près des **principes FAIR** (Facile à trouver, accessible, interopérable et réutilisable).



Qu'est-ce qu'Omeka ?

Développé aux États-Unis en 2008 par le Center for History and New Media (CHNM) de l'Université George Mason (également à l'origine du logiciel de gestion bibliographique Zotero), Omeka est arrivé en France aux alentours de 2010.

Au départ, il a été conçu pour mettre en ligne des musées virtuels, mais il s'est déployé dans d'autres domaines. À partir de données de recherche brutes, l'outil permet de créer des **bases de données éditorialisées**, autrement dit structurées, accessibles, et visibles sur le web. L'outil offre une grande modularité des fonctionnalités grâce à de nombreux plugins, et traite les divers objets multimédia (textes, images, sons, vidéos).

Quels avantages pour les établissements de recherche ?

Au-delà de répondre aux principes FAIR en affichant des **contenus interopérables**, les établissements de recherche qui exposent leurs données de recherche via Omeka s'inscrivent dans la démarche de **science ouverte** fortement encouragée par le CNRS. De plus, l'outil offre plusieurs avantages techniques :

- l'interface est simple et intuitive ;
- les métadonnées sont moissonnables, permettant notamment le référencement dans d'autres bases ;
- une base Omeka peut être connectée à d'autres services grâce à une API REST.

Des sites conçus avec les utilisateurs

La technique utilisée pour Omeka est maîtrisée à l'Inist, l'institut s'étant pourvu de ce logiciel en 2015. L'équipe de l'Inist chargée de l'exposition des données a répondu à sa première demande externe en 2016, plus précisément au Centre Camille Jullian (UMR CNRS/Aix-Marseille Université). C'est en étroite collaboration qu'ils ont créé la plateforme Corpus & Ressources Archéologiques (CoReA) qui a pour objectif de valoriser des fonds d'archives multimédias et des carnets de terrain en archéologie. Le site web CoReA a vu le jour en 2017, suivi du site IMAGE en 2018 avec le laboratoire HisCant-MA (Histoire et cultures de l'antiquité et du Moyen Âge (Université de Lorraine).

Aujourd'hui, plusieurs sites construits avec Omeka sont en ligne et une dizaine de projets sont en cours de mise en œuvre :

- Institut des mondes africains – IMAF (CNRS, Université Panthéon-Sorbonne, EHESS, IRD, Université Aix-Marseille, EPHE) ;
- Histoire et cultures de l'antiquité et du Moyen Âge (Université de Lorraine).

Ces sites ont été réalisés via une **démarche UX**, une méthode de conception et d'évaluation de l'expérience utilisateur. Reformatage des données, enrichissement, alignement, géolocalisation ou encore conception graphique sont autant d'actions qui ont été accomplies en étroite collaboration avec les établissements à l'origine de la demande.

De plus, une fois la base de données mise en ligne, des agents de l'Inist forment le personnel du laboratoire chargé de gérer le nouveau site. Celui-ci étant hébergé à l'Inist, l'institut offre un suivi et une assistance.

Un développement en vue

Avec l'objectif de développer ce service, l'équipe, aujourd'hui constituée de 14 agents, travaille sur le déploiement de sites web avec Omeka. Le but est de proposer des versions plus riches en fonctionnalités et selon des présentations différentes. Ce façonnage en amont permettra de gagner en efficacité pour répondre aux demandes de mise en ligne de bases de données éditorialisées.

Lodex : quand la science rencontre la culture

Publié le 25 mai 2020

Généralement utilisé pour structurer les données issues de travaux de recherche, le logiciel open source Lodex, produit par l'Inist, a été utilisé par des étudiants pour structurer les données issues de la bibliographie de l'exposition « Le merveilleux scientifique ».

Tout a commencé par une sollicitation émanant d'un enseignant de l'Ecole nationale supérieure des sciences de l'information (Enssib) en novembre 2019. En effet, après avoir lu l'article « LODEX : des données structurées au web sémantique » paru sur HAL*, l'enseignant a contacté l'Inist pour une collaboration autour d'un travail sur le traitement et la visualisation des données.

Plusieurs étapes

La première étape a consisté à aller à la rencontre des étudiants pour leur présenter Lodex dont la fonction est de faciliter la curation et la sémantisation de données brutes pour les connecter au web de données via les normes et les standards du web sémantique. Face aux 19 étudiants en Master 2 – Politique des bibliothèques et de la documentation, les intervenantes de l'Inist, Anne-Marie Badolato et Stéphanie Gregorio, ont commencé par une phase plutôt théorique afin que les étudiants découvrent l'outil. Ces derniers ont ensuite eu plusieurs semaines pour se l'approprier avant que le binôme Inist y retourne pour une nouvelle formation axée davantage sur la pratique.



Quatre résultats pour une même source

Pour mettre en pratique l'utilisation de Lodex dans le cadre d'un devoir, les étudiants ont travaillé sur la bibliographie de l'exposition « Le merveilleux scientifique – une science-fiction à la française » présentée en 2019 à la BnF. Ils se sont divisés en quatre groupes de travail, ce qui a permis de mettre en évidence qu'à partir d'une même source, les résultats peuvent être différents en fonction de l'utilisation de Lodex :

- le premier groupe a présenté un résultat très descriptif avec un affichage de textes riches (https://enssib-groupe1.lodex.inist.fr/) ;
- le deuxième groupe s'est orienté vers un affichage créatif mais plus pauvre (https://enssib-groupe2.lodex.inist.fr/) ;
- le troisième groupe a exposé un grand nombre de données mais moins fouillées (https://enssib-groupe3.lodex.inist.fr/) ;
- le quatrième groupe a affiché des notices et des données enrichies avec des liens vers Wikidata et DataBNF (https://enssib-groupe4.lodex.inist.fr/).

Lodex, à mettre dans toutes les mains

Cette utilisation de Lodex, réalisée pour la première fois avec des étudiants, a prouvé que cet outil pouvait être approprié par des usagers divers, avec des objectifs différents. De plus, leur retour a été unanime : LODEX est un outil relativement simple à utiliser et clé en main.

Suite à cette expérience qui s'est couronnée de succès, les équipes de l'Inist sont prêtes à répondre à d'autres demandes pour faire découvrir Lodex.

*Stéphanie Gregorio, Alain Collignon, François Parmentier, Nicolas Thouvenin. LODEX : des données structurées au web sémantique. *Atelier Web des Données de la 19ème Conférence sur l'Extraction et la Gestion des Connaissances (EGC 2019)*, Jan 2019, Metz, France. (hal-01990444).



OPIDoR Tour, retour sur une belle aventure

Publié le 1er octobre 2020

Pendant un an, les équipes de l'Inist sont allées à la rencontre des chercheurs, enseignants-chercheurs, professionnels de l'IST et chargés de valorisation pour leur présenter les outils et services OPIDoR.

Cet **OPIDoR Tour**, on l'a imaginé, on l'a fait ! **L'heure est au bilan**…

Pourquoi un OPIDoR Tour ?

Depuis 2019, l'Agence nationale de la recherche (ANR), dans sa feuille de route inscrite dans le plan national de la Science ouverte, demande que les projets financés élaborent et intègrent un plan de gestion des données (DMP) dans les six mois suivant le démarrage dudit projet. Ce document synthétise la description et l'évolution des jeux de données, il prépare le partage, la réutilisation et la pérennisation des données.

Etant spécialisé dans la gestion, le partage et la diffusion des données de la recherche, c'est tout naturellement que l'Inist s'est proposé pour aller présenter les outils et services OPIDoR (DMP OPIDoR ainsi que les services associés Cat OPIDoR et PID OPIDoR) aux quatre coins de la France.

Du travail d'équipe

Les agents du service OPIDoR et ceux du service Formation ont mis leurs compétences en commun pour apporter une réelle expertise en matière de gestion de données.

Sur place, chaque étape s'est inscrite dans une démarche concertée avec les acteurs locaux et régionaux. Les sessions ont été organisées en liens étroits avec les délégations régionales du CNRS et les services de proximité des universités. **Nous remercions à nouveau nos collègues de terrain pour leur précieuse collaboration** notamment au niveau de la logistique, pour les inscriptions aux sessions ou encore pour l'accueil des participants. Des intervenants locaux ont également apporté des compléments personnalisés.

OPIDoR Tour en chiffres

Au total, **19 sessions** ont été organisées pour un nombre global de 703 participants. 43 % d'entre eux ont répondu à l'enquête de satisfaction :

- **74 %** = taux global de satisfaction avec un record pour les ateliers GopenDoRe (84 %) suivi par la présentation des outils (77 %).
- **60 %** des répondants ne connaissaient pas les outils OPIDoR et **83 %** sont prêts à les utiliser suite à la présentation.
- **17** intervenants des services OPIDoR et DoRANum + **19** intervenants en région (IPE, documentalistes, cellules données de la recherche…) ont animé les sessions.

Le bilan

Même si les déplacements pour l'OPIDoR Tour ont été semés d'embuches (grèves, neige… et Covid !), les agents de l'Inist ayant participé à cette aventure ont particulièrement apprécié **le contact** avec leurs interlocuteurs et usagers. Ils ont pu prendre le temps d'échanger et de mieux identifier les besoins. De leur côté, les participants ont salué le déplacement de l'Inist sur le terrain et le fait de pouvoir identifier les personnes pouvant répondre à leurs questions.

Cette aventure a aussi permis de conforter l'Inist dans sa mission d'accompagnement et de conseil autour de la gestion des données, une activité majeure au sein de l'un des trois axes forts de l'Institut « Valoriser les données de la recherche ».

Face à ce succès, l'Inist pourrait organiser d'autres « tours » sur d'autres thématiques.

À suivre…

Évaluation « FAIRitude » pour Loterre et TermSciences

Publié le 19 octore 2020

Au départ : une sollicitation de l'Inrae pour tester une grille d'évaluation FAIR. À l'arrivée : une étude qui a permis de démontrer que Loterre répond en grande partie aux principes FAIR.

FAIR, vous avez dit FAIR ?

Les principes FAIR sont un ensemble de principes directeurs pour gérer les données de la recherche visant à les rendre **faciles à trouver, accessibles, interopérables et réutilisables** par l'homme et la machine. Ces principes s'inscrivent dans le cadre de la science ouverte qui vise à faciliter l'accès aux publications scientifiques et aux données de la recherche.

Le projet FooSIN pour développer la culture FAIR

Répondant à un appel Flash science ouverte de l'ANR, le projet FooSIN (ANR-19-DATA-0019) est la participation française au GO FAIR Food Systems Implementation Network. Autour de l'agriculture et de l'alimentation, l'objectif du projet est notamment de proposer des outils et des méthodologies pour « **FAIRifier** » les données et ressources sémantiques.

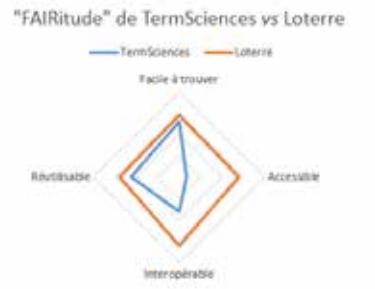
C'est dans le cadre du développement de ces outils que l'Inrae, partenaire du projet FooSIN, a sollicité l'Inist en juin 2020 pour tester une grille d'évaluation : la grille SHARC développée par RDA (Research Data Alliance) dont l'Inist fait partie.

Loterre et TermSciences comme cas d'usage

Le service Ingénierie terminologique de l'Inist s'est donc prêté au jeu qui consistait à évaluer, selon les critères proposés par la grille SHARC, la prise en compte des principes FAIR par les plateformes d'exposition de terminologies TermSciences et Loterre, toutes deux conçues par l'Inist à des époques bien différentes.

En effet, quinze ans séparent la mise en œuvre des deux plateformes et l'analyse comparée des résultats permet de prendre la mesure de l'évolution des pratiques d'exposition de ressources sémantiques à l'Inist.

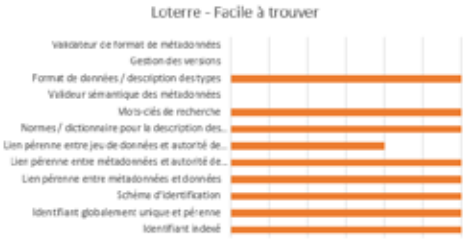
Dans cette étude « test », un score a été appliqué aux différents champs de la grille. Une note globale a ainsi été attribuée à chaque critère FAIR ; des chiffres utilisés dans l'élaboration de schémas pour une lecture plus visuelle des résultats. Le radar de comparaison (ci-dessous) en est une illustration.



TermSciences contre Loterre

Sans surprise en raison des années qui séparent les deux plateformes, Loterre s'est révélée plus « FAIR » que TermSciences. Suivant l'évolution générale du monde de la recherche scientifique, la plateforme Loterre (lancée en 2018) est résolument tournée vers la **science ouverte** et la prise en compte des principes de partage et de réutilisation des données.

Exemple de critère : Loterre – Facile à trouver



Outre le constat, cette analyse a également permis d'entrevoir les axes qui pourraient permettre à Loterre d'améliorer encore sa « FAIRitude ».

Une voie que l'équipe de l'Inist compte bien suivre…

Lire aussi…

Romain David, Laurence Mabile, Alison Specht, Sarah Stryeck, Mogens Thomsen, et al.. FAIRness Literacy: The Achilles' Heel of Applying FAIR Principles. *CODATA Data Science Journal, Committee on Data for Science and Technology (CODATA)*, 2020, 19 (32), pp.1–11. (10.5334/dsj-2020–032) https://dx.doi.org/10.5334/dsj-2020–032

ISTEX-DL pour un accès facilité aux publications scientifiques

Publié le 4 novembre 2020

Toujours dans le but de faciliter l'accès aux publications scientifiques, l'application ISTEX-DL (ISTEX DOWNLOAD), spécialisée dans l'extraction de corpus, a bénéficié d'améliorations afin de répondre aux demandes des utilisateurs.

Mieux exploiter l'archive ISTEX

ISTEX, c'est l'accès à 23 millions de documents de littérature scientifique dans toutes les disciplines, pour la communauté française de l'enseignement supérieur et de la recherche. Au milieu de cette masse de documents, comment élaborer des requêtes et télécharger simplement… sans être informaticien ?

C'est pour répondre à cette question que l'application ISTEX-DL a été créée en 2017.

En effet, jusqu'alors, extraire un corpus de documents ISTEX supposait une certaine pratique de l'informatique, les solutions proposées étant l'utilisation de moissonneurs – utilitaires en ligne de commande – ou l'emploi de la fonction « extract », fonction du moteur de recherche de l'API ISTEX se paramétrant directement dans l'URL d'interrogation.

La création de l'application ISTEX-DL a permis à tous les ayants droit d'exploiter l'archive ISTEX pour en extraire un corpus de documents adapté à ses besoins, et ce sans connaissance informatique particulière.

Découvrez la vidéo de présentation d'ISTEX-DL sur la chaine YouTube de l'unité

Quoi de neuf dans cette version ?

L'application ISTEX-DL a évolué vers une interface plus ergonomique, plus proche de la charte ISTEX, et elle est dotée de nouvelles fonctionnalités.

La principale nouveauté est le dépassement des limites dans la taille des requêtes simples ou complexes (plus aucune limite !) et dans le nombre de documents à télécharger (jusqu'à 100 000 documents !).

Parmi les autres nouvelles fonctionnalités proposées :

- Choix du nombre de documents et du mode de tri (pertinence et qualité) ;
- Affichage d'un échantillon du corpus afin de vérifier les premiers résultats ;
- Usage personnalisé des formats d'export : choix dans les exports possibles en fonction des usages et téléchargement des données au format directement compatible avec vos outils de traitements ou de visualisation ;
- Extraction rapide de gros volumes ;
- Possibilité de partager le corpus.

Ces nouvelles fonctionnalités et les modifications apportées avec cette version sont expliquées de façon détaillées dans le blog Istex : ISTEX-DL se met sur son 4.21 et sort le grand jeu

Pour faire quoi ?

ISTEX-DL permet ainsi de télécharger facilement du texte intégral dans l'archive ISTEX. Il est aussi possible de récupérer l'ensemble de sa bibliographie, de faire une recherche documentaire dans les archives et d'obtenir tous les documents, d'**extraire des gros volumes de documents pour des applications de text mining**…

À chacun son utilisation en fonction de ses besoins.

Prochainement, des tutoriels ISTEX-DL

Et pour vous aider avec la prise en main de l'application et découvrir en détail, toutes les fonctionnalités, **des tutoriels sont en cours de réalisation**, mise en ligne prochainement dans notre rubrique dédiée.

La nouvelle version d'ISTEX-DL est en ligne : https://dl.istex.fr/

N'hésitez pas à faire part de vos retours, www.istex.fr/contact ou via la liste users (users @ listes.istex.fr sans espace). L'équipe ISTEX est à votre disposition pour répondre à toute question.



La veille… plus que jamais d'actualité !

Publié le 19 novembre 2020

L'information scientifique et technique (IST) est un domaine en constante évolution. À l'heure où la recherche avance à grands pas, il est primordial d'accompagner le monde de la recherche et les professionnels du domaine en mettant à disposition une mine d'informations actualisées. C'est ce que propose l'Inist via son service LaLIST, véritable base de connaissances, et aussi au travers de la veille opérée pour le site Ouvrir la science piloté par le Comité pour la science ouverte (CoSO).

LaLIST

Ouvert en 2012, le site LaLIST a fortement évolué en huit ans. Son ambition : donner aux lecteurs des informations fiables, validées, et leur faire gagner du temps.

Avec quelque 10 000 billets publiés, il est aujourd'hui une référence parmi les sites d'actualités en information scientifique et technique (IST) abondamment citée.

Toute une méthodologie

Deux agents de l'Inist, Brigitte Clément et Isabelle Didelot, alimentent LaLIST. Toutes deux ont recensé plusieurs centaines de sites nationaux et internationaux ainsi que des réseaux sociaux pouvant être sources d'informations à partager.

Mais la veille, au-delà de l'abondance d'informations, c'est aussi un cadre : il faut détecter et cibler l'information, utiliser les bons outils, et s'assurer de la fiabilité des éléments à qui vont être diffusés.

La ligne éditoriale de LaLIST s'appuie sur le projet Inist dans lequel s'inscrivent les trois axes forts de l'Institut : Accéder à l'information scientifique, Valoriser les données de la recherche, Analyser et fouiller l'information scientifique. Parmi les thématiques phares, on retrouve les données de la recherche, le TDM (text and data mining) et tous les autres enjeux de la Science ouverte dans le monde entier.

Chaque jour, des centaines d'informations sont passés en revue via divers outils d'alerte, de suivi de newsletters institutionnelles ou autres, de personnalités et d'experts, des flux RSS et une surveillance des réseaux sociaux. De là sont extraits des articles qui collent à l'actualité et aux thématiques de LaLIST. Plusieurs champs sont ensuite repris pour le partage : titre /URL/source/auteur/date de l'article/mots-clés/extrait significatif de l'article.

Une base de connaissances

LaLIST compte environ **3100 abonnés** à ce jour issus majoritairement de l'enseignement supérieur et de la recherche en France et à l'étranger. Six nouveaux billets (français et anglais) sont publiés chaque jour ouvré, toute l'année. En plus d'être un site d'actualité, LaLIST est une véritable base de connaissances puisqu'il comporte un **moteur de recherche**, une grande valeur ajoutée.

LaLIST, c'est un site web, mais c'est aussi une page Facebook et un compte Twitter. Une newsletter est également diffusée de façon quotidienne ou hebdomadaire (selon le choix de l'abonné).

Au vu des « like » sur les réseaux sociaux et des messages envoyés par les utilisateurs, nul doute que LaLIST a su convaincre les professionnels de l'IST sur l'intérêt du service en matière de veille.

La revue *Arabesques* a d'ailleurs sollicité l'Inist pour un article intitulé « LaLIST, veille et diffusion de l'information scientifique et technique » publié dans le cadre d'un dossier dédié à la veille*.

L'équipe LaLIST est accompagnée par le service Web de l'Inist pour la partie technique et pour l'ajout de nouvelles fonctionnalités liées entre autres au partage.

Un projet de refonte du site est actuellement à l'étude.

Site LaLIST : https://lalist.inist.fr/

*Lire : "LaLIST, veille et diffusion de l'information scientifique et technique" Brigitte Clément, article publié dans *Arabesques* janvier-février-mars 2019, p. 10–11 : https://publications-prairial.fr/arabesques/index.php?id=102

Ouvrir la science

Piloté par le Comité pour la science ouverte (CoSO), le site Ouvrir la science est entièrement dédié à la **science ouverte**.

Dans la continuité des sites d'actualité sur le mouvement du libre accès

La création du site Ouvrir la science a commencé en avril 2018 avec le concours du service Web de l'Inist et de Thérèse Hameau, chargée de veille à l'Institut. Elle suit le mouvement du libre accès depuis le début des années 2000. Elle a été la cheville ouvrière du site Libre accès, lancé en 2003, et de celui sur les données de la recherche, créé à l'initiative du ministère de l'Enseignement supérieur et de la Recherche (MESR) en 2012. Si ces sites ont été arrêtés en janvier 2019, son activité éditoriale et de veille se poursuit à travers le site Ouvrir la science, lancé à l'occasion des Premières Journées nationales de la science ouverte qui se sont tenues en novembre 2018.

Le site Ouvrir la science s'inscrit dans la continuité de ces deux sites et du site BSN (Bibliothèque scientifique numérique) du MESR, et se tourne résolument vers la science ouverte.

Ses objectifs sont de trois ordres : valoriser le travail du Comité pour la science ouverte et ses productions ; fournir des informations, des documents, des outils pour la science ouverte principalement aux chercheurs et personnels d'appui à la recherche ; exprimer les orientations politiques de la France en matière de science ouverte et relayer celles des autres pays en Europe et dans le monde.

Le site bénéficie d'une version anglaise à laquelle collaborent les membres du service traduction de l'Inist.

Tous les champs de la science ouverte

Le travail de veille opéré à l'Inist s'articule notamment autour du **blog d'actualité** ainsi que des rubriques « **Ressources** » et « **Agenda** ».

Il s'appuie sur le suivi de plus de 300 sites, sur des listes de diffusion, des lettres d'information, des flux RSS… En moyenne, un article est publié par semaine et 5 ressources par mois.

Soutenu par le ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation (MESRI), le site affiche une sensibilité politique : il prend en compte les domaines législatifs ainsi que les directives européennes et les textes officiels.

Une référence

Avec environ **900 abonnés** à la newsletter mensuelle, le site Ouvrir la science est devenu une référence en matière de science ouverte en raison de la pertinence des informations et des documents signalés.

Le comité de rédaction, constitué par des agents de l'Inist et par des représentants du MESRI, passe en revue les différentes rubriques chaque mois et élabore la newsletter.

Le site s'accompagne d'un compte Twitter suivi par plus de **2 700 abonnés**.

Site Ouvrir la science :
https://www.ouvrirlascience.fr/

TÉMOIGNAGE

« La veille, c'est un métier mais c'est aussi un état d'esprit d'ouverture et de curiosité. Il faut avoir envie d'apprendre et de transmettre. »

Brigitte Clément
chargée de veille à l'Inist



Le service Publications participe à la nouvelle collection du MNHN

Publié le 17 décembre 2020

Le Muséum national d'histoire naturelle (MNHN) vient de lancer sa nouvelle collection Natures en Sociétés avec la parution des trois premiers volumes réalisés en collaboration avec le service Publications de l'Inist.

Retour sur cette collaboration

Depuis plusieurs années, un travail de secrétariat d'édition est réalisé à l'Inist en partenariat avec le MNHN pour la revue Anthropozoologica qui est désormais publiée en flux continu.

Fort de ses compétences en matière de stylage et dans l'utilisation de la chaîne éditoriale Métopes, le service Publications a été sollicité par le Muséum pour sa nouvelle collection Natures en sociétés prévue en numérique et en version papier.

L'Inist avait notamment en charge la préparation des manuscrits dans la chaîne Métopes, les contrôles biblio-typo-ortho-grammaticaux, le traitement des images quand nécessaire ou encore la cohérence des contenus.

Covid oblige, la collaboration a dû se faire à distance, entre échanges de mails et réunions en visio, mais le contexte n'a pas entaché la bonne mise en œuvre du projet.

Des échanges à distance

Démarré en janvier 2020, le projet a commencé par une phase test sur un volume et s'est poursuivi par le stylage et la constitution des index sur les deux autres volumes, l'incorporation des images, l'intégration des corrections d'auteurs, le renvoi des fichiers et la mise au point des XML (structuration CSS et corrections html) pour les e-pub.

L'équipe Inist a également suivi une formation Métopes avec le Muséum en visio, dispensée par le groupe de formation Métopes de l'Université de Caen, afin de calibrer les volumes dans XMLMind et de modifier la CSS ou le HTML.

Une parution en deux temps

Les trois premières parutions concernent dans un premier temps les versions papier et EPUB. La mise en ligne de ces trois volumes se fera dans un deuxième temps sur la plateforme OpenEdition Books.

Cette collection en sciences humaines et sociales s'intéresse aux relations entre les sociétés, contemporaines ou passées, et leurs natures.


Retrouvez ce rapport d'activité en ligne sur son site dédié
cnrs Inist Institut de l'information scientifique et technique
https://ra2020.inist.fr


À bientôt !

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