



Histoire de la recherche contemporaine

La revue du Comité pour l'histoire du CNRS

Tome X - n°2 | 2021

Les outils documentaires de la science

Du papier à la science ouverte : évolutions des services d'accès à l'information

From paper to open science : developments in information access services.

Arnaud Moign, Cécilia Fabry, Anne-Marie Badolato et Claire François



Édition électronique

URL : <https://journals.openedition.org/hrc/6735>

DOI : 10.4000/hrc.6735

ISSN : 2265-786X

Éditeur

CNRS Éditions

Référence électronique

Arnaud Moign, Cécilia Fabry, Anne-Marie Badolato et Claire François, « Du papier à la science ouverte : évolutions des services d'accès à l'information », *Histoire de la recherche contemporaine* [En ligne], Tome X - n°2 | 2021, mis en ligne le 31 décembre 2021, consulté le 20 novembre 2022. URL : <http://journals.openedition.org/hrc/6735> ; DOI : <https://doi.org/10.4000/hrc.6735>

Ce document a été généré automatiquement le 20 novembre 2022.

Tous droits réservés

Du papier à la science ouverte : évolutions des services d'accès à l'information

From paper to open science : developments in information access services.

Arnaud Moign, Cécilia Fabry, Anne-Marie Badolato et Claire François

Cet article est le résultat d'entretiens réalisés par Arnaud Moign avec plusieurs personnes de l'Inist qui ont imaginés, développés, fait vivre et promus ses services. Il en résulte forcément une certaine hétérogénéité de traitement, des déséquilibres, des imprécisions. Un grand merci à Laurence Grand, Brigitte Clément, Patrick Kremer, Dolorès Dardaine, Jacqueline Gillet, Catherine Fournier, Christine Weil-Miko, Dominique Lechaudel, Stéphane Gully, Nicolas Thouvenin, Jean-François Nominé et Thérèse Hameau.

- 1 Retracer l'évolution des services d'accès à l'information scientifique ces 30 dernières années permet de mesurer l'impact du développement de l'informatique, d'internet avec une capacité de diffusion en ligne toujours plus rapide et efficace.
- 2 Dans les années 1980 -1990, l'édition sous forme papier est la règle, les premières bases de données sont accessibles sur le minitel. Cet article présente ici l'organisation mise en place pour fournir des copies d'articles à la demande principale alternative à l'abonnement aux revues sous forme papier. Il montre l'adaptation progressive des services qui prennent en compte les opportunités, apportées par la diffusion des documents directement sous forme numérique, avec l'ouverture des portails documentaires. Enfin, il décrit comment ces évolutions ont permis l'essor de la notion de Science Ouverte et la mise en place de politiques favorisant sa mise en œuvre. Les professionnels de l'IST, rompus aux techniques de documentation d'une production scientifique, ont toute leur place pour relever ce nouveau défi.

L'âge d'or du papier

La création de l'Inist

- 3 L'Institut de l'information scientifique et technique (Inist) a été créé en 1988, suite à la fusion des deux centres de documentation historiques du CNRS, le Centre de documentation scientifique et technique et le Centre de documentation en sciences humaines. Sa création s'est faite dans un contexte politique de redéploiement économique (effets de la crise sidérurgique dans le bassin lorrain) avec la délocalisation des activités documentaires du CNRS et la volonté de former un unique centre capable de fournir l'accès à la documentation scientifique de manière moderne.



Site de l'INIST

D.R.

- 4 Cette nouvelle entité a donc été pensée comme une « usine », une ruche très active - accueillant plus de 400 personnes au plus fort de son activité - dont la mission était d'assurer l'accès à l'information pour le CNRS, les autres organismes de recherche publique, et également pour les entreprises privées, avec des modèles économiques adaptés aux différentes cibles. Les deux activités principales sont la fourniture de copie d'articles scientifiques et la production des bases de données PASCAL et FRANCIS.

A cette époque et jusqu'en 2015, l'Inist produit deux bases de données signalant les documents publiés dans la plupart des domaines des sciences et techniques

PASCAL

Base de données bibliographiques multidisciplinaire et multilingue qui couvre l'essentiel de la littérature mondiale en science, technologie et médecine.

Cette base contient plus de 20 millions de références bibliographiques depuis 1973. Elle est reconnue comme base de première approche pour toute recherche scientifique et technique

FRANCIS

Base de données bibliographiques multidisciplinaire et multilingue en sciences humaines, sciences sociales et économiques qui contient 2.6 millions de références depuis 1972.

L'industrialisation de la fourniture « papier »

- 5 L'activité de fourniture de document primaire (FDP) consiste à procurer la copie d'un document scientifique (article, thèse, rapport, etc.) en réponse à la demande des différents utilisateurs du service. L'industrialisation de cette activité est fortement liée à la volonté de mieux répondre à la demande et gagner en productivité. L'application informatique de gestion développée au début des années 1990 permet ainsi l'enregistrement de l'ensemble des commandes et le suivi de chaque étape du traitement.
- 6 En 1994, par souci d'efficacité, l'organisation du service est repensée et la notion de « ligne de produit » est intégrée, en vue de remplacer les différentes équipes indépendantes. L'idée est d'obtenir un suivi précis des chaînes de commande pour gagner en rapidité. Le service est alors constitué de quatre lignes de produits : deux pour le service courant, une pour le service rapide et une dernière dédiée aux traitements spéciaux, un service à valeur ajoutée d'identification de références bibliographiques et de localisation du document dans une bibliothèque extérieure. Chaque ligne de produit est constituée d'un(e) secrétaire commercial(e) en relation avec les utilisateurs, de documentalistes qui enregistrent les commandes, de magasiniers qui préparent les documents à photocopier et de personnes chargées de l'expédition des documents. Au plus fort de l'activité, à la fin des années 1990, le nombre de commandes traitées dans une année dépasse les 700 000 et l'ensemble du service compte plus de 50 personnes.

Un service performant et international

- 7 Pendant des années, l'Inist demeure le premier fournisseur français de copie de documents scientifiques, pour le public comme pour le privé. Cette image positive de référent est due en grande partie aux efforts déployés par ses nombreux documentalistes, dans la recherche documentaire. En effet, lorsque la demande d'un chercheur ne peut être satisfaite grâce aux collections disponibles dans le fonds documentaire – plusieurs millions de documents stockés sur 32 km de rayonnages ou sur disques optiques numériques : 27 000 titres de revues et plusieurs dizaines de milliers de rapports scientifiques, comptes rendus de congrès ou thèses au milieu des années 1990 – le service recours prend le relais. Créé en 1997, sa raison d'être est de tout mettre en œuvre pour trouver le document demandé et de le commander auprès d'un réseau de bibliothèques françaises ou étrangères aussi prestigieuses que la British Library. A cette époque, ce service regroupe une quinzaine de documentalistes, la part

de commandes honorées grâce à ces bibliothèques dépasse 20 % au début des années 2000.

- 8 En procédant de cette manière, l'Inist satisfait près de 98 % des commandes qui lui sont faites.

Le tournant des années 2000 : l'arrivée du numérique et le déclin progressif du papier

- 9 Le début des années 2000 marque un tournant dans l'histoire de ce type de service intermédiaire entre l'utilisateur et l'éditeur, et ce pour deux raisons.
- 10 La première concerne le développement fulgurant d'internet et la proposition par les éditeurs de services d'accès directs et payants à leurs collections, selon deux modèles : l'abonnement à des bouquets entiers de revues ou le principe de pay-per-view, un accès limité et payant à l'article. La deuxième raison est liée à l'explosion à la fois du nombre de revues scientifiques et de leur coût qui limitent le développement des fonds documentaires.
- 11 En 2002, la décroissance du nombre de commandes s'amorce et cette tendance se confirme au fil des ans. Alors qu'en 2001 l'Inist fournit 682 000 copies d'articles par an, ce nombre est réduit à seulement 158 000 en 2012 et 6 800 à la fermeture du service de fourniture de documents en 2018.

La fourniture de documents électroniques et la création de RefDoc

- 12 Pour répondre aux changements dans les pratiques des utilisateurs, plusieurs évolutions voient le jour comme la livraison en format PDF directement sur le poste de travail du chercheur. Ce nouveau service de fourniture de documents électroniques sécurisée par un DRM (Digital Rights Management, système de gestion des droits numériques) nécessite des accords pour l'accès à environ 2 000 revues : des licences permettant de fournir des articles électroniques à partir des collections en ligne des éditeurs et des autorisations pour numériser à la demande des articles à partir des collections Inist.
- 13 Plus tard, en 2008, un nouvel outil permettant aux utilisateurs de disposer d'une interface à guichet unique pour la recherche de références et la commande de copie de documents est proposé. À son lancement RefDoc¹ offre 52 millions de références d'articles en ligne. Comme sur n'importe quel site de e-commerce, l'utilisateur peut alors paramétrer sa commande, choisir le nombre de copies, le délai de livraison, suivre sa commande à tout moment et choisir son mode de paiement : carte bancaire, PayPal pour les utilisateurs qui ne souhaitent pas créer de compte.
- 14 Parallèlement à ces développements, l'Inist met progressivement en place un accès dématérialisé à l'information scientifique. Pourtant, l'arrivée des premiers portails thématiques - permettant aux chercheurs de consulter en ligne des articles de revues - n'a pas signé l'arrêt immédiat de la fourniture de copie de document. En réalité, la transition vers le tout numérique a été extrêmement douce, puisque le service RefDoc n'a pris fin officiellement qu'en 2018, même si en 2013 il n'est déjà plus accessible qu'aux organismes publics.

L'ère des portails documentaires²

- 15 A la fin des années 1990, l'ensemble du personnel de l'Inist a engagé une réflexion sur l'avenir de ses services et en particulier sur l'impact du développement du document électronique.

ConnectSciences : le premier portail générique

- 16 Suite à cette réflexion, la création d'une première « application électronique » qui permet à des internautes d'accéder à de l'information en ligne est actée, il s'agit de ConnectSciences.
- 17 ConnectSciences offre l'accès à plusieurs centaines d'URLS classées par thèmes scientifiques. Les utilisateurs peuvent alors retrouver des références d'articles, en aucun cas le texte intégral de ces documents, le cas échéant, ils peuvent ensuite passer commande à l'Inist.
- 18 Ce portail qui s'adresse à l'ensemble de la communauté scientifique a été le socle de tous les autres car sa création a nécessité de répondre à un certain nombre de questions fondamentales autour de l'authentification des personnes, des droits accordés aux utilisateurs, mais aussi autour du circuit de commande. Ce premier portail ne connaît cependant pas un très grand succès car il ne donne accès qu'à l'information librement accessible ce qui est encore très rare pour les publications scientifiques. Cette première expérience a néanmoins eu le mérite d'éveiller la curiosité de la communauté scientifique.

Les portails thématiques

- 19 Après cette première version « très ouverte », l'Inist décide d'offrir l'accès au texte intégral des documents scientifiques en proposant le portail multidisciplinaire BiblioSciences, dès 2001, qui existe en parallèle de ConnectSciences.
- 20 Cette décision implique la mise en place de négociations avec les différents éditeurs pour des abonnements permettant aux chercheurs du CNRS d'accéder en ligne aux documents nécessaires à leur activité de recherche.
- 21 Ainsi BiblioSciences propose un ensemble de ressources (accès aux articles via les plates-formes Science Direct, Springer Link, ou encore accès à la revue Nature et aux bases de données du Web of Knowledge, aux bases de données via la plate-forme Ovid SP, Engineering Village) aux près de 32 000 chercheurs, ingénieurs et techniciens du CNRS.
- 22 Le premier portail thématique créé n'est pas un portail CNRS, puisqu'il s'agit de BiblioInserm, un portail ouvert en 2002 à la demande de l'Inserm qui veut faire évoluer sa structure d'IST, mais ne dispose pas des ressources informatiques nécessaires. Ce portail dédié aux chercheurs de l'Inserm propose les ressources de plusieurs éditeurs, négociées conjointement par l'Inist et l'Inserm. Presque deux ans après sa création, le succès de ce premier portail thématique conduit à la création du premier portail thématique CNRS destiné aux chercheurs en biologie : BiblioVie. Pendant plusieurs années, ces deux portails vivent ainsi côte à côte, avec des ressources négociées en commun.

- 23 Les réalisations commencent à susciter de plus en plus d'intérêt de la part des différents départements du CNRS, et certains désirent disposer de leur propre portail. Aussi sont successivement créés TitaneSciences (Chimie) en 2003, BiblioSHS (Sciences humaines et sociales, arts, lettres) en 2004 et BiblioST2I (sciences et technologies de l'information et la communication) en 2006.
- 24 En 2015, l'Inist gère 7 portails différents qui offrent l'accès à plus de 12 000 revues électroniques, 15 000 livres électroniques, et 43 bases de données.

Vers la création d'un portail unique : BibCnrs

- 25 Les évolutions technologiques et les différentes campagnes de tests et audits d'ergonomie régulièrement menés auprès des chercheurs ont conduit les équipes à opérer une simplification et une modernisation régulière de ces différents portails. En 2016, une nouvelle refonte de ces derniers est décidée, avec cette fois la volonté de créer un portail unique, BibCnrs.
- 26 La première version de ce portail propose une interface très orientée « communautés » avec un accès par instituts ; la version actuelle propose l'accès aux ressources au sein d'une page unique avec 3 onglets : un article, une revue - un ouvrage, une base de données. Cette évolution est le résultat d'une démarche UX qui utilise différentes approches : enquêtes, entretiens, observations de pratiques des utilisateurs/testeurs.
- 27 Les ressources documentaires auxquelles peuvent accéder par les différents portails, d'abord chacun des 6 départements, puis chacun des 10 instituts du CNRS, sont toujours négociées avec les éditeurs en fonction des besoins de chaque structure (fig.1).

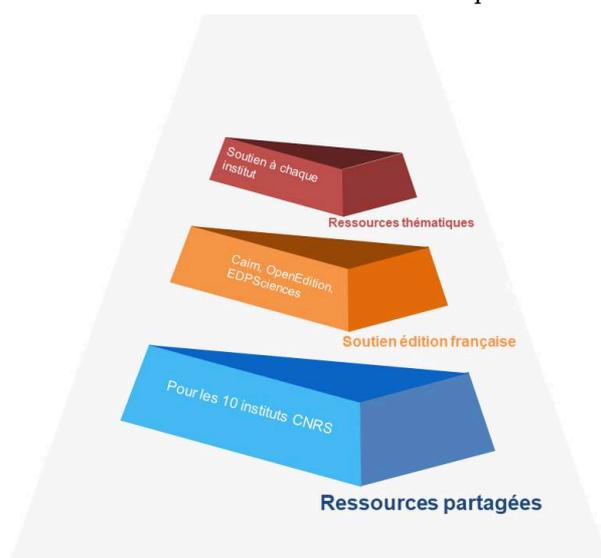


Figure 1 -Les ressources accessibles via BibCnrs

- 28 En parallèle de l'évolution de BibCnrs, l'Inist développe d'autres services, afin de diversifier les usages des ressources scientifiques. L'un des meilleurs exemples est probablement celui de la plate-forme ISTEEX qui permet la fouille de texte.

La plate-forme ISTEEX : pour un développement de la fouille de texte

- 29 La plate-forme ISTEEX est une gigantesque archive multidisciplinaire contenant plus de 23 millions de publications numériques, provenant de 30 corpus de littérature scientifique, rédigées en plus de 50 langues et couvrant une période extrêmement large, qui s'étend du XIV^e siècle à nos jours.
- 30 ISTEEX est un projet porté par le CNRS et piloté par un ensemble de partenaires : le consortium Couperin, l'Agence bibliographique de l'Enseignement supérieur (ABES) et l'Université de Lorraine, dans le cadre des Initiatives d'excellence du programme « Investissements d'Avenir » initié par le ministère de l'Enseignement supérieur et de la Recherche. Ce projet financé par l'Agence nationale de la recherche (ANR) à hauteur de 60 M€ permet ainsi la construction de la plate-forme ISTEEX, une tâche qui est confiée à l'Inist.³
- 31 L'un des objectifs d'ISTEEX est d'amorcer un changement de modèle pour l'accès aux ressources scientifiques électroniques proposées par les éditeurs en « *passant d'une logique de flux à une logique de stockage* », ceci afin de ne pas dépendre d'un accès dont la pérennité n'est pas assurée.
- 32 La plateforme ISTEEX est pleinement opérationnelle et reconnue par les chercheurs depuis plusieurs années maintenant. Néanmoins, l'utilité d'ISTEEX ne s'arrête pas au stockage de document puisque cette plate-forme a été pensée dès le départ pour permettre une exploitation poussée des données contenues dans ses documents, à savoir la fouille de texte ou le « Text and Data Mining (TDM) ». Les statistiques d'ISTEEX pour l'année 2020 – près de 5 millions de consultations pour environ 350 établissements français de l'enseignement supérieur et de la recherche (ESR) – montrent que la motivation de la consultation est également répartie entre usage documentaire et fouille de texte.
- 33 Cette facette d'ISTEEX est beaucoup plus technique et nécessite l'utilisation d'outils informatiques puissants. Pour favoriser le développement de cet usage, les documents disponibles via ISTEEX sont enrichis de catégories scientifiques, d'une indexation, d'une annotation des entités nommées ou par la structuration des références citées⁴. Des corpus thématiques spécialisés sont régulièrement créés à partir des documents ISTEEX et proposés aux chercheurs du domaine concerné⁵.
- 34 Pour aider les chercheurs dans leur exploitation des corpus extraits de la plate-forme, les équipes travaillent actuellement au développement de solutions de visualisation des données dont différents types de cartographies.

L'accès aux ressources documentaires : une perpétuelle évolution

- 35 Les outils numériques évoluent sans cesse, de même que les pratiques des chercheurs. Conscient de cet état de fait, l'Inist fait évoluer ses outils afin qu'ils soient toujours plus performants et ergonomiques.
- 36 Ces dernières années, les équipes ont préparé plusieurs « plug-in » que les chercheurs ont pu intégrer dans leur environnement de travail pour accéder à l'information disponible sur les différentes plates-formes comme BibCnrs ou ISTEEX.
- 37 L'extension Click & Read propose désormais aux chercheurs d'accéder directement au texte intégral des plates-formes, via une icône qui apparaît au cours d'une recherche,

qu'elle soit effectuée sur un moteur de recherche, un site web ou directement sur le site d'un éditeur.

L'adaptation des pratiques : mutualisation et mesure des usages

Les négociations auprès des éditeurs : une tâche complexe

- 38 Au début des années 2000, la dématérialisation des ressources documentaires et la création des premiers portails ont poussé tous les acteurs IST vers de nouvelles pratiques dans les négociations avec les éditeurs. Ces négociations concernent aussi bien la nature des abonnements, que le nombre d'utilisateurs, la question du droit d'auteur ou même la pérennité des accès aux ressources déjà acquises.
- 39 *Négocier la pérennité des accès : une nécessité*
- 40 Lorsque l'édition papier était « la norme », l'achat d'une ressource permettait une consultation permanente et sans limite de temps, notamment grâce aux archives physiques. Cette pérennité est remise en cause avec l'arrivée du numérique, les éditeurs ayant un contrôle total sur l'accès aux ressources électroniques.
- 41 Cet aspect est alors intégré dans les négociations avec les éditeurs et conduit à la mise en place de PANIST, un serveur qui garantit un accès aux archives des ressources négociées par le consortium Couperin⁶. Ce serveur géré, maintenu et hébergé par l'Inist offre ainsi aux établissements qui sont totalement désabonnés un accès permanent aux archives. Néanmoins, à la différence d'ISTEX il ne permet pas la fouille de texte.
- 42 *Poids des grands éditeurs et négociations communes*
- 43 Au fil des ans, les compétences en négociation de l'Inist se sont élargies. En effet, les achats de documentation et abonnements gérés par les départements et instituts du CNRS ont progressivement été regroupés au niveau de l'Inist.
- 44 À l'image d'autres organismes de recherche comme l'Inrae ou l'Institut Pasteur, le CNRS (via l'Inist) s'est longtemps chargé seul des négociations avec les éditeurs. En parallèle, les organismes plus petits, conscients de leur faible impact se sont regroupés autour du Consortium Couperin afin de mutualiser les négociations au niveau national. Le CNRS a finalement intégré Couperin et en 2005 le premier groupement de commandes national intégrant entre autres le CNRS, l'Inserm et Couperin a été signé.
- 45 Aujourd'hui l'acquisition des ressources se fait à plusieurs niveaux : si une partie des contenus est toujours négociée par l'Inist en direct pour les laboratoires du CNRS, une autre part est gérée de manière commune, au niveau national, par le consortium Couperin.

La mesure des usages : une tâche stratégique

- 46 La mesure des usages est essentielle pour accompagner la politique documentaire d'un établissement. L'arrivée du numérique a demandé la mise en place de nouvelles méthodes de comptage. L'Inist s'est ainsi intéressé à ces questions dès le début des années 2000.

ezPAARSE et ezMESURE : les outils de comptage

- 47 Pour mesurer les usages de ressources qui sont proposées sur les plates-formes documentaires, l'Inist s'appuie aujourd'hui sur 2 outils : ezPAARSE et ezMESURE.
- 48 ezPAARSE⁷ est un logiciel qui identifie en temps réel et de façon anonyme les consultations de chaque communauté de chercheurs sur une plate-forme comme BibCNRS, en respectant les recommandations Counter⁸. Les informations qu'il fournit permettent de faire le lien entre l'usage réel des ressources électroniques et le prix des abonnements.
- 49 Ce logiciel libre développé par l'Inist en 2013 est un outil de niche très spécialisé et installé dans de nombreux établissements en France mais aussi à l'étranger. C'est le cas par exemple avec le consortium Online Computer Library Center (OCLC) qui a intégré ezPAARSE sur ses propres plates-formes.
- 50 Créé en 2016, ezMESURE, est une plate-forme centralisée qui permet de recueillir l'ensemble des informations de consultation de chaque entité ezPAARSE et de générer des rapports personnalisés proposant des graphiques à partir des événements de consultation. Comme pour ezPAARSE, le développement est réalisé en partenariat avec Couperin, grâce à un financement du ministère de l'Enseignement supérieur, de la Recherche et de l'innovation (MESRI).
- 51 Le partage et l'analyse croisée des mesures des usages des plates-formes documentaires des différents établissements sont importants pour harmoniser et optimiser les politiques d'achat et de négociations des ressources et définir une politique documentaire nationale.
- 52 L'augmentation régulière des coûts d'abonnement aux ressources documentaires s'est perpétuée voire accentuée avec l'accès en ligne. En lien ou en réaction à ses phénomènes le mouvement pour le libre accès puis celui de la science ouverte amorce une nouvelle révolution.

Les perspectives apportées par le développement des politiques de science ouverte

Du mouvement Open Access à la science ouverte

- 53 Le mouvement Open Access qui prône la libre diffusion des savoirs scientifiques trouve son origine au début des années 2000, alors que l'internet est en train de transformer en profondeur les réalités matérielles et économiques de la diffusion de la connaissance scientifique. En quelques années, plusieurs événements vont définir les contours et objectifs de l'Open Access. Ainsi, l'initiative de Budapest pour l'accès ouvert (BOAI) de 2002 est rapidement suivie de la déclaration de Bethesda pour l'édition en libre accès en 2003. Cette même année, la déclaration de Berlin sur le libre accès à la connaissance⁹ élargit la notion du libre accès à « toutes les œuvres et données, issues d'un travail de recherche, à tous les champs disciplinaires et au patrimoine culturel ». Ces textes sont considérés à ce jour comme les 3 textes fondateurs du mouvement Open Access.
- 54 Depuis plusieurs années, les pouvoirs publics à l'échelle française et européenne se sont emparés de ces questions. En France, l'article 30 de la loi pour une république numérique autorise depuis 2016 les chercheurs à déposer leurs publications en libre

accès, si elles sont issues d'une activité de recherche financée par plus de 50 % de fonds publics. Ce droit est dévolu aux chercheurs après une période d'embargo de 6 à 12 mois, pour n'importe quel contrat établi entre le chercheur et l'éditeur de la revue. Par ailleurs, la Commission européenne va jusqu'à rendre obligatoire le libre accès aux publications issues de recherches ayant reçu des financements du Conseil européen de la Recherche (ERC) ou via le programme Horizon 2020, sous peine de sanctions financières.

- 55 En 2018, le MESRI lance le Plan national pour la science ouverte (PNSO), avec pour objectif de mettre en place les conditions du développement d'une science ouverte en France. Ce premier plan 2018-2021 a conduit entre autres à la création du Comité pour la science ouverte (CoSO). Parmi les réalisations de ce plan se trouvent : l'adoption d'un nouveau modèle de financement et de gouvernance pour HAL, le soutien à des modèles économiques d'édition en accès ouvert et sans frais de publications pour les auteurs (modèle diamant), la formalisation d'un projet de plate-forme nationale sur les données de la recherche : *Recherche Data Gouv*. Les réflexions du CoSO et l'ensemble des réalisations peuvent être suivies sur le site Ouvrir la Science¹⁰.
- 56 Le deuxième Plan¹¹, annoncé en juillet 2021 est encore plus ambitieux. Avec un budget de 15 M€ par an pour la science ouverte, il se donne pour objectifs de généraliser les pratiques de science ouverte en France et de structurer, partager et ouvrir les données de la recherche. Ce deuxième PNSO est ainsi en complète cohérence avec la loi de programmation de la recherche de 2020 qui fixe un objectif de 100 % de publications en accès ouvert en 2030. L'inscription des pratiques de science ouverte dans la durée nécessite également de faire évoluer le système d'évaluation.

Le rôle de l'Inist dans le contexte Open Access

- 57 Très tôt l'Inist s'intéresse à ce mouvement : en janvier 2003, au Carré des Sciences à Paris est organisée la première journée sur le libre accès¹²¹³. En tant que co-organisateur, l'Inist joue alors un rôle de premier plan en lançant un site d'information - libre accès à l'information scientifique et technique¹⁴ - qui signale les informations concernant les différentes initiatives des acteurs de l'information scientifique. L'Inist est également à l'origine de la signature de la Déclaration de Berlin par le CNRS.
- 58 Néanmoins, s'il est indéniable que des initiatives internes ont été prises très tôt, les services proposés alors ne sont pas encore en accord avec ces principes.
- 59 Les idées ont lentement infusé et ces dernières années, l'implication de l'Inist dans la science ouverte s'est fortement développée. Cette réflexion a finalement entraîné l'arrêt de la FDP et de la production des bases de données PASCAL et FRANCIS, activités historiques, et le développement de nouvelles offres en accord avec les problématiques de la science ouverte, à savoir des services concernant les données de la recherche ou la possibilité d'utiliser des publications scientifiques à des fins d'analyse.

Les données de la recherche : une nouvelle mission

- 60 Depuis 2015, la mission d'IST du CNRS, définie dans le décret n° 82-993 du 24 novembre 1982¹⁵ a ainsi été revue et son champ d'action s'est élargi. Il ne s'agit plus seulement de documentation et de publication ; Le CNRS gère désormais officiellement les données de la recherche et des plates-formes documentaires.

- 61 Dans la version modifiée par le décret n° 2015-1151 du 16 septembre 2015, l'article 2 stipule ainsi que le CNRS doit :
- 62 « Assurer l'élaboration et la diffusion de la documentation scientifique et la publication des travaux et **données de la recherche**, notamment en mettant à disposition de la communauté scientifique et universitaire des plateformes documentaires et en contribuant à leur enrichissement. »

Outiller les chercheurs pour faciliter la gestion et le partage des données

- 63 Mettre en place une véritable politique de science ouverte est un défi qui nécessite d'accompagner les chercheurs en leur proposant des services à la fois fonctionnels et accessibles.
- 64 Dans ce cadre, l'Inist a développé la plate-forme DoRANum¹⁶, dont le but est d'accompagner la communauté scientifique dans son appropriation de la démarche de gestion et de partage des données. Cette plate-forme qui rassemble toutes sortes de contenus de formation est entièrement constituée de ressources libres sous forme de vidéos, de fiches synthétiques, de tutoriels, de quizz.

La plate-forme OPIDoR et la gestion des données de la recherche

- 65 Collecter et gérer les données de la recherche est une tâche complexe, qui nécessite une réflexion profonde en amont. L'Inist est impliqué depuis longtemps dans ces sujets : la participation à des groupes de travail de la Research Data Alliance (RDA) est effective depuis 2014 et est à l'origine du développement de la plate-forme OPIDoR¹⁷.
- 66 Cette dernière donne accès à un catalogue des services français dans la gestion des données, propose un support pour la rédaction des plans de gestion de données (DMP) et pour attribuer des d'identifiants pérennes (DOI). Au cœur de l'accompagnement des communautés scientifiques, elle est pleinement opérationnelle et reconnue : le modèle de plan de gestion de données de l'ANR étant dorénavant accessible et utilisable via DMP OPIDoR.
- 67 Par ailleurs, une plate-forme nationale fédérée de données de la recherche, appelée Recherche Data Gouv¹⁸ sera disponible à la fin du premier trimestre 2022. L'Inist participera à ce projet d'entrepôt national multi-domaines piloté par l'Inrae comme fournisseur de services d'accompagnement notamment via les plates-formes OPIDoR et DoRANum.
- 68 Cette initiative constitue une vraie opportunité pour l'Inist, qui trouve ainsi toute sa place dans l'organisation nationale autour des données de la recherche.



Inist-CNRS
D.R.

Conclusion

- 69 En 30 ans, le paysage de l'information scientifique a subi une mutation considérable, la révolution du numérique a frappé de plein fouet une organisation bien rôdée où chacun, chercheurs, éditeurs, intermédiaires et bibliothèques apportait sa pierre à l'édifice.
- 70 L'informatisation des processus et la normalisation des métadonnées (les formats MARC partagés internationalement) ont largement facilité le travail de chacun permettant une efficacité croissante, ce qui est décrit sous le terme d'« industrialisation ». La capacité de diffuser directement les textes sous forme numérique a totalement rebattu les cartes apportant un confort indéniable aux chercheurs qui obtiennent leur article directement sur leur poste de travail, mais obligeant l'ensemble des acteurs à revoir leurs pratiques pour s'adapter aux nouveaux besoins. C'est ce que l'Inist a fait en développant les portails numériques et en arrêtant la fourniture de copies d'articles et la production des bases de données bibliographiques PASCAL et FRANCIS.
- 71 Cette évolution s'accompagne de la diffusion et du partage des différentes productions scientifiques aussi variées que les supports pédagogiques, les logiciels et bien sûr les données de recherche, c'est le socle des politiques de science ouverte déployées au niveau international. Le défi pour les acteurs de l'IST est de participer activement au développement de l'environnement favorable facilitant les bonnes pratiques. Les besoins sont considérables, de nouveaux métiers (comme les *data stewards* ou *data analysts*) se dessinent. L'Inist s'inscrit dans cette dynamique en développant les services d'accompagnement à la gestion, au partage et à l'analyse des productions scientifiques y compris les données de recherche.

NOTES

1. Laurence Grand, « Refdoc à l'INIST-CNRS. La référence en fourniture de documents », *Arabesques* [En ligne], 63 | 2011, mis en ligne le 25 juin 2020, consulté le 22 juillet 2021. URL : <https://publications-prairial.fr/arabesques/index.php?id=1850>
2. Drouard Françoise, Petitjean Sophie, Schöpfel Joachim, « Les portails d'information du CNRS », *Communication & langages*, 2009/3 (N° 161), p. 75-92. DOI : 10.4074/S033615000900307X. URL : <https://www.cairn.info/revue-communication-et-langages1-2009-3-page-75.htm>
3. Présentation du projet ISTEEX. ministère de l'Enseignement supérieur et de la recherche. 2012. http://www.cpu.fr/wp-content/uploads/2012/04/Projet_istex.pdf
4. Métriques concernant les enrichissements ISTEEX. <http://enrichments-metrics.data.istex.fr/>
5. Camille de Salabert, Sabine Barreaux. Vers un corpus optimal pour la fouille de textes : stratégie de constitution de corpus spécialisés à partir d'ISTEX. 6^e conférence conjointe Journées d'études sur la parole (JEP, 33^e édition), Traitement automatique des langues naturelles (TALN, 27^e édition), Rencontre des étudiants chercheurs en informatique pour le traitement automatique des langues (RÉCITAL, 22^e édition). Volume 4 : Démonstrations et résumés d'articles internationaux, Jun 2020, Nancy, France. pp. 66-69. (hal-02768520v3) <https://hal.archives-ouvertes.fr/hal-02768520v3>
6. Présentation de Couperin. Historique. <https://www.couperin.org/presentation/historique>
7. L'outil open source ezPAARSE est disponible gratuitement ici : <https://github.com/ezPAARSE-project/ezPAARSE>
8. Projet Counter <https://www.projectcounter.org/>
9. Déclaration de Berlin sur le libre accès à la connaissance en sciences exactes, sciences de la vie, sciences humaines et sociales. 22 octobre 2003 https://www.ouvrirelascience.fr/wp-content/uploads/2018/11/D%C3%A9claration-Berlin_fr.pdf
10. Ouvrir la science <https://www.ouvrirelascience.fr/>
11. Deuxième Plan national pour la science ouverte. Généraliser la science ouverte en France 2021-2024 <https://www.ouvrirelascience.fr/deuxieme-plan-national-pour-la-science-ouverte/>
12. Libre accès à l'information scientifique et technique : État de l'art et perspectives - janvier 2003 <https://openaccess.inist.fr/libre-acces-a-linformation-scientifique-et-technique-etat-de-lart-et-perspectives-janvier-2003/>
13. Battisti Michèle, « Libre accès à l'information scientifique et technique : état de l'art et perspectives », *Documentaliste-Sciences de l'Information*, 2003/1 (Vol. 40), p. 37-45. DOI : 10.3917/docs.401.0037. <https://www.cairn.info/revue-documentaliste-sciences-de-l-information-2003-1-page-37.htm>
14. Libre accès à l'information scientifique et technique <https://openaccess.inist.fr/>
15. Décret n° 82-993 du 24 novembre 1982 portant organisation et fonctionnement du Centre national de la recherche scientifique. <https://www.legifrance.gouv.fr/loda/id/JORFTEXT00000884328/>
16. Le portail des données de la recherche apprentissage numérique <https://doranum.fr/>
17. OPIDoR <https://opidor.fr/>
18. Recherche Data Gouv : plate-forme nationale fédérée des données de la recherche. <https://www.ouvrirelascience.fr/recherche-data-gouv-plateforme-nationale-federee-des-donnees-de-la-recherche/>

RÉSUMÉS

Cet article présente l'évolution de l'accès à l'information scientifique, au travers des services proposés par l'Inist, depuis la fourniture à la demande de copie papier d'articles jusqu'à la mise à disposition du texte intégral sous forme électronique pour la fouille de texte. Parmi les étapes et transformations importantes sont présentées la mise en place des portails documentaires, son impact sur les négociations avec les éditeurs scientifiques pour la définition des périmètres de l'offre, la montée en puissance de l'accès ouvert ainsi que la mesure des usages permettant de suivre l'évolution des pratiques des utilisateurs. Enfin, cet article décrit comment cette évolution implique la diffusion et le partage des données de recherche, dans le cadre du déploiement des politiques de Science Ouverte.

This article presents the evolution of access to scientific information as reflected in the range of services provided by the Inist - from the supply of paper copies of articles on request to the provision of full text in electronic form for text mining. The significant stages and transformations of that evolution include the creation of document portals, the impact of this on negotiations with scientific publishers to define the scope of the offer, the rise in importance of open access as well as usage measurement which monitors changes to user practices. Finally, this article describes how this evolution necessarily involves the dissemination and sharing of research data in the framework of the implementation of Open Science policies.

INDEX

Mots-clés : Accès à l'information - Fourniture de document - Portail documentaire - Négociations - Mesure des usages - Fouille de texte - Données de recherche - Science Ouverte

Keywords : Access to information - Document supply - Document portal - Negotiations - Usage measurement - Text mining - Research data - Open Science

AUTEURS

ARNAUD MOIGN

Rédacteur indépendant et ingénieur, pigiste depuis 2017. Diplômé de l'École nationale supérieure d'ingénieurs de Limoges, diplômé de recherche technologiques (DRT), a exercé 10 ans en tant qu'ingénieur spécialisé en traitements de surface et en analyse du cycle de vie (ACV), au sein du CITRA à Limoges.

CÉCILIA FABRY

Responsable communication veille de l'Inist-CNRS, responsable du service communication, veille & web de l'Inist depuis 2015. De formation scientifique, elle a tout d'abord occupé les fonctions de chargée de veille pendant plus de 10 ans avant de travailler comme chef de projet statistiques d'usages dans le cadre du projet ezPAARSE.

ANNE-MARIE BADOLATO

Chargée de veille en science ouverte au sein de l'Inist. Elle a précédemment occupé des fonctions de responsabilité dans les activités de fourniture de documents ou d'appui au pilotage

scientifique. Entre 2000 et 2008 elle est chargée de veille et de prospective pour les questions liées au libre accès.

CLAIRE FRANÇOIS

Directrice de l'Inist-CNRS, docteur en science du sol et titulaire d'un DESS en informatique, elle rejoint l'Inist en 1990. Elle y a porté une activité de R&D dans le domaine de l'analyse et la fouille de l'information scientifique et successivement pris la responsabilité de départements déployant l'offre de service de l'Inist. Depuis septembre 2018, elle assure la direction de l'Inist.
claire.francois@inist.fr