



HAL
open science

ISIDORE : une plateforme de recherche de documents et d'information pour les Sciences Humaines et Sociales

Stéphane Pouyllau

► **To cite this version:**

Stéphane Pouyllau. ISIDORE : une plateforme de recherche de documents et d'information pour les Sciences Humaines et Sociales. 2011. <sic_00605642>

HAL Id: sic_00605642

https://archivesic.ccsd.cnrs.fr/sic_00605642v1

Preprint submitted on 3 Jul 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

ISIDORE : Une plateforme de recherche d'information pour les Sciences Humaines et Sociales¹

ISIDORE est un nouvel outil au service des scientifiques pour rechercher des informations dans les données numériques des sciences humaines et sociales. Créé par le TGE Adonis, ISIDORE est accessible principalement via le site web www.rechercheisidore.fr et offre la possibilité de naviguer dans plus d'un million de documents provenant de 860 sources (mai 2011) : revues, ouvrages, bases de données bibliographiques et scientifiques, carnets de recherche (blogs), entrepôts numériques de données pour la recherche, actualités scientifiques, bibliothèques universitaires et municipales. Ce nombre de sources est appelé à augmenter régulièrement.

Qu'est-ce qu'ISIDORE ? A qui s'adresse-t-il ? En quoi ISIDORE est-il différent des autres moteurs de recherche ? Comment l'utiliser ? Comment diffuser ses données dans ISIDORE ? Réponses...

Une plateforme de recherche innovante

Depuis quelques années, la recherche en sciences humaines et sociales (SHS) utilise et produit massivement des données numériques. Les sources utilisées pour l'exercice de la recherche, les séminaires virtuels, les publications en ligne, sont autant d'exemples qui montrent que des pans entiers du processus d'élaboration du savoir utilisent des données nativement numériques ou résultant d'un processus de numérisation.

Ces données, mais aussi les outils de traitement et d'édition numériques, ont petit à petit transformé le bureau de l'enseignant-chercheur, de l'ingénieur et du doctorant, en un bureau numérique qui est également « en ligne ». Parallèlement à ce mouvement, l'utilisation du web comme support d'« une base de données mondiale », où les données sont liées entre-elles, s'accélère. Ce web de données, appelé aussi « web 3.0 » porté par le W3C (consortium international gérant le web) et un grand nombre de projets de recherche universitaire, sera sans doute la prochaine évolution du web.

ISIDORE est le premier outil de cette génération. Son rôle est de rendre accessibles, via un point d'accès unifié, les productions numériques des SHS. Il valorise également le travail et les méthodes des professionnels de l'information scientifique et technique qui classent, sélectionnent, numérisent, structurent, diffusent et archivent les données numériques : une chaîne de production qui s'étend des archives aux revues électroniques.

ISIDORE est une plateforme informatique qui regroupe plusieurs outils :

- un collecteur d'informations moissonnant les données selon des standards et protocoles informatiques internationaux ;
- une chaîne de traitement de l'information permettant d'enrichir les données en qualifiant l'information collectée à l'aide de référentiels scientifiques (de thésaurus par exemple) ;
- un moteur de recherche offrant un accès à cette information enrichie ;
- une interface web pour le public : c'est le site rechercheisidore.fr ;
- un entrepôt de données RDF permettant l'accès à l'information enrichie selon les méthodes du web de données : c'est le point d'accès et d'interrogation en langage SPARQL.

A qui s'adresse ISIDORE ?

¹ L'auteur remercie Nadine Dardenne (TGE Adonis), Jean-Luc Minel (TGE Adonis), Shadia Kilouchi (TGE Adonis), Sophie David (TGE Adonis) ainsi que la direction du TGE Adonis : Jean-Luc Pinol et Richard Walter pour les relectures de ce document.

Pour les étudiants, le site rechercheisidore.fr permet d'explorer les disciplines scientifiques, et de se constituer un complément bibliographique de ressources en ligne. Au doctorant, il offre la possibilité de suivre des questions scientifiques très pointues en interrogeant simultanément plusieurs revues électroniques, des archives ouvertes et des bases de données bibliographiques en seulement quelques clics. Pour l'enseignant, c'est un outil pertinent pour enrichir ses cours de références scientifiques actuelles. Pour l'enseignant-chercheur, en tant que lieu d'exposition des données de la recherche, ISIDORE est une plateforme particulièrement indispensable tout au long du processus de recherche : se référer à des références scientifiques validées ou critiquées, mettre à jour une bibliographie, explorer un champ disciplinaire proche du sien, s'informer sur les sujets de thèses, faire de la veille... Pour les bibliothécaires et les documentalistes, ISIDORE, est un outil de veille informationnelle. Il offre la possibilité de transformer une requête en flux de veille scientifique et l'utilisation des facettes permet d'affiner les réponses ou de ré-ouvrir vers de nouvelles questions.

Que trouve-t-on dans ISIDORE ?

ISIDORE collecte des informations (métadonnées), des données et des actualités venant d'institutions des SHS, nous les appelons « producteurs ». La plupart du temps, les producteurs de données font acte de candidature et le TGE Adonis analyse les conditions scientifiques et techniques du moissonnage. Il s'agit de plateformes de revues, de bibliothèques de recherche, de centres de documentation, d'équipes de recherche fonctionnant en projets, de laboratoires qui ont fait l'effort - depuis parfois plusieurs années - de rendre leurs données interopérables selon des standards et normes internationales. ISIDORE contient en majorité des données en accès libre (ce que nous privilégions). Pour autant, les données sous barrière mobile et dont l'accès est contrôlé, ne sont pas exclues à la condition qu'une ou plusieurs solutions d'accès à ces données existent pour les communautés scientifiques. Ainsi, nous signalons tous les articles des revues présentes sur la plateforme Cairn, y compris ceux dont le texte intégral n'est pas accessible gratuitement : l'accès pouvant se faire pour l'instant soit depuis la plateforme biblioshs.inist.fr pour les unités mixtes de recherche, soit via certains environnements numériques de travail dans les universités et les bibliothèques universitaires.

ISIDORE contient des informations enrichies par croisement avec des référentiels métiers (thésaurus, vocabulaires contrôlés, référentiels) produits par les scientifiques ou par des institutions de la recherche, de l'enseignement et de la culture. Par exemple, la bibliothèque numérique *Gallica* de la Bibliothèque nationale de France, la bibliothèque universitaire de Rennes 2, le Centre de documentation REGARDS du laboratoire ADES du CNRS ou encore la base de données *Horizon-Pleins Textes* de l'Institut de Recherche sur le Développement (IRD)..

ISIDORE permet de rechercher en une seule requête dans les publications scientifiques éditées par Revues.org, Persée, ou Cairn mais aussi dans le contenu des articles déposés dans HALSHS réalisé par le Centre pour la communication scientifique directe. Il donne également accès aux thèses déposées dans TEL ou dans le futur portail theses.fr réalisé par l'ABES. Les nouvelles formes de publications scientifiques sont présentes : les carnets de recherche d'Hypotheses.org, très riches, sont indexés par ISIDORE. Les photographies scientifiques déposées dans MédiHAL sont aussi présentes. Enfin, les actualités scientifiques venant de Calenda.org, de Fabula.org, et plusieurs dizaines de sites web de laboratoires et de Maisons des Sciences de l'Homme sont accessibles et interrogeables au travers d'ISIDORE.

En quoi ISIDORE est-il différent des autres moteurs de recherche ?

ISIDORE se distingue des moteurs de recherche classiques, car les données qu'il expose proviennent d'institutions académiques et scientifiques : elles ont donc une réelle valeur scientifique. ISIDORE ne collecte et n'indexe pas directement le web, il ne contient que ce que les producteurs lui ont indiqué de prendre ; ainsi dans ISIDORE, il n'y a pas de pages web de type « A propos » ou « menu », il n'y a pas non plus de publicité et les réponses sont tout de suite pertinentes. Pour reprendre un exemple de l'une des étudiantes ayant testé ISIDORE lors de sa phase de test : à la requête « garnis », seules les références scientifiques sur les hôtels meublés des XIXe et XXe sont affichées. Bouquets et autres paniers « garnis », sans intérêt pour cette recherche, ne viennent pas polluer les résultats.

Comment l'utiliser ?

ISIDORE est accessible via un site web, www.rechercheisidore.fr. A partir de la page d'accueil, à l'ergonomie simple, vous avez accès à un véritable tableau de bord qui vous offre une multitude de fonctions : champs de recherche, différentes facettes permettant la navigation dans les données, filtrage des résultats, outils de visualisation des enrichissements, possibilité de rouvrir le spectre d'une recherche ou de rebondir, bouton permettant d'enregistrer une requête sous la forme d'un flux de veille spécifique (RSS), signalement d'une ressource par email ou sur les plateformes de réseaux sociaux (facebook, twitter, etc.) etc. Cette interface riche, simple et claire, vous permet d'utiliser toute la puissance d'ISIDORE et d'être très rapidement autonome. Certaines bibliothèques universitaires proposent déjà le site web www.rechercheisidore.fr dans leur portail documentaire, n'hésitez pas solliciter vos documentalistes et bibliothécaires pour avoir des conseils d'utilisation. Très prochainement des tutoriels vidéo viendront illustrer le potentiel de la plateforme.

Comment diffuser ses données dans ISIDORE ?

ISIDORE valorise à la fois les sources pour la recherche (archives scientifiques, archives de chercheurs, de laboratoires) et les publications. Dans ce cadre, l'ensemble des équipes peuvent participer à l'alimentation d'ISIDORE. Pour cela, plusieurs voies sont possibles :

- Les revues : Si votre revue est hébergée sur Revues.org, Cairn, Persée ou I-Revues, elles sont déjà présentes dans ISIDORE car nous avons associé ces éditeurs à la réalisation d'ISIDORE. Si votre revue n'est pas hébergée par l'un de ces portails, il existe deux façons de la connecter à ISIDORE : l'actualité de vos numéros peut être relayée sur ISIDORE via les flux d'actualité de type RSS (voir l'encadré « Une question/une réponse ») ou mieux, le signalement de tous vos articles, classés par numéros, peut être récupéré par ISIDORE si votre outil d'édition possède un export OAI-PMH des métadonnées. ISIDORE « moissonnera » - c'est le terme consacré - les métadonnées et indexera vos articles. De plus, si vos articles sont en HTML, PDF, MSWord ou OpenOffice, le texte intégral sera également moissonné, si vous l'autorisez.

- Les bases de données bibliographiques : ISIDORE peut moissonner les bases de données bibliographiques même si, dans un premier temps, le TGE Adonis privilégie les bases de données diffusant des données en texte intégral, des images, ou des données multimédia. Pour cela, comme pour les revues, ISIDORE utilise le protocole OAI-PMH, massivement utilisé dans les communautés scientifiques, pour moissonner les données.

- Les actualités scientifiques : ISIDORE collecte et indexe l'ensemble des événements scientifiques signalés dans Calenda.org et plusieurs dizaines de flux d'actualité (RSS) venant des laboratoires de recherche et des projets d'équipes. N'hésitez pas à contacter le TGE Adonis pour signaler le flux

d'actualité de votre laboratoire ou de votre projet.

ISIDORE est également un projet innovant : c'est le plus gros projet d' « *open data* » scientifique en France. Ainsi, il est capable de moissonner et d'indexer finement les informations contenues dans les données scientifiques. Pour cela, il suffit de structurer l'information d'une page web selon le modèle de description des données RDF (pour *Resource Description Framework*). Déjà plusieurs producteurs d'information, présents dans ISIDORE, utilisent RDF pour structurer leurs données.

ISIDORE offre ainsi un point d'accès à l'information scientifique et technique. Conçu et développé par le TGE Adonis, il est ouvert à tous les scientifiques. Sa réalisation a été possible grâce à un partenariat avec le [Centre pour la communication scientifique directe](#), qui a assuré la maîtrise d'œuvre, et des marchés de réalisation avec les sociétés Atos Consulting, Antidot, Sword et Mondéca. Sa réalisation a pris un peu plus d'une année entre l'écriture du cahier des charges (précédé d'une phase de définition débutée en 2007) et la mise en ligne de la version de test public (version bêta) lors de la [deuxième Université](#) du TGE Adonis en décembre 2010.

Actuellement, ISIDORE est en version 1. Cependant, comme tous les outils de recherche d'information, il devra évoluer et être alimenté en permanence (on n'imagine pas une plateforme de recherche figeant ses évolutions) car il est à la fois un catalyseur de données et de publications qui valorise et rend visible l'information scientifique et un espace interdisciplinaire où chacun peut explorer les autres disciplines. C'est pourquoi il est important de maintenir les actions de recherche et développement, d'expérimentation autour de cet outil : les communautés scientifiques doivent utiliser, critiquer, et s'approprier celui-ci pour une visibilité croissante des activités et des productions en sciences humaines et sociales.