

scienceplus.abes.fr : une nouvelle base de données au service de la science ouverte



Lancée en mars 2021, scienceplus.abes.fr est la toute dernière réalisation du projet mené par l'Abes pour l'enrichissement et la mise à disposition des métadonnées des ressources numériques spécifiques à l'ESR. Elle apporte une réponse à de multiples besoins exprimés par les établissements. Sa valeur ajoutée est de fournir un signalement au niveau des articles et des chapitres, sous une forme harmonisée, enrichie, et interopérable.

La base de données scienceplus.abes.fr¹ expose les métadonnées de corpus de publications scientifiques numériques (revues, ebooks, articles, chapitres) pour lesquelles l'Abes a obtenu le droit de réutilisation de la part des éditeurs, notamment dans le cadre des programmes d'achat de documentation sous licence nationale (ISTEX et CollEx-Persée). Dans le Sudoc, ces métadonnées sont intégrées au niveau du titre (revue, ebook). La valeur ajoutée de scienceplus.abes.fr réside principalement dans le signalement au niveau des articles et des chapitres, sous une forme harmonisée, enrichie et interopérable.

Précisons tout d'abord que le service scienceplus.abes.fr n'a pas vocation à devenir un grand index de centaines de millions de publications scientifiques, qui viserait l'exhaustivité comme les services offerts par *Google Scholar*, *Dimensions*, *lens.org*, *Central Discovery Index*. En revanche, il vise à combler certaines lacunes du périmètre de bases géantes ou à aller plus en finesse quand c'est nécessaire et souhaité (focus sur les chercheurs français, indexation avec des vocabulaires francophones, prise en charge de métadonnées de gestion, non bibliographiques, périmètre précis des licences nationales françaises).

Le Github de l'Abes² présente le tableau qui liste et décrit les corpus chargés. Ces informations sont également mentionnées en RDF dans la base elle-même. À ce jour, celle-ci contient plus de 6 millions d'articles.

Interopérabilité et Web de données

scienceplus.abes.fr est une base RDF interrogeable en SPARQL³, conforme

aux principes et mécanismes du *Linked Open Data* (LOD). En versant ces métadonnées sur le LOD, l'Abes contribue à la construction d'un Web de données au service de la science ouverte. Les technologies du Web sémantique offrent en effet une grande flexibilité, ce qui facilite l'exploration et la réutilisation des métadonnées, sans préjuger des besoins des utilisateurs.

Pour autant, ces technologies ne constituent pas la seule manière de promouvoir l'interopérabilité, et la base scienceplus pourra à l'avenir étoffer son offre de services pour s'adapter au contexte de travail des réutilisateurs, en premier lieu les bibliothèques de l'ESR. Le support du protocole OAI-PMH constitue une priorité.

Des données enrichies

Outre l'**harmonisation et le reformatage des métadonnées** fournies par les éditeurs, scienceplus.abes.fr propose des **enrichissements** significatifs, principalement sur trois points :

- Les mentions d'auteur sont alignées sur le **référentiel IdRef (540 000)** et le **registre Orcid (440 000)**. Cette identification favorise l'exploration des données et leur interopérabilité à l'échelle nationale et globale. Les documents liés à IdRef enrichissent également idref.fr et son triple store data.idref.fr.
- La base scienceplus.abes.fr extrait et met en valeur selon les principes et mécanismes du LOD, l'indexation sujet des éditeurs basée sur des vocabulaires contrôlés (MSC, JEL, etc.).
- Les métadonnées des publications achetées sous licences nationales sont ouvertes, mais, d'un contrat à l'autre, cette ouverture n'a pas la même portée.

Les conditions de réutilisation sont explicitement précisées dans le tableau listant les corpus et parmi les triplets eux-mêmes.

Un service Bêta à l'écoute des besoins de l'ESR

Le service scienceplus.abes.fr évoluera en fonction des besoins exprimés par la communauté française de l'ESR. La priorité est donnée aux métadonnées des corpus peu ou mal signalés dans les outils commerciaux. Depuis le lancement du service en mars 2021, l'Abes rencontre les équipes de bibliothèques pour mieux appréhender les contenus actuels et les possibilités ou contraintes de leurs outils. Au moment de la rédaction de cet article (6 septembre 2021), le versement des métadonnées de la totalité des articles et chapitres d'un important agrégateur est en préparation.

Chacun est invité à s'exprimer *via* le guichet d'assistance⁴ afin d'indiquer aux équipes de l'Abes quels contenus leur seraient utiles, quelles données modéliser/enrichir/connecter/diffuser et pour répondre à quels besoins.

YANN NICOLAS

Responsable du Labo de l'Abes
nicolas@abes.fr



[1] <https://scienceplus.abes.fr>
[2] <https://github.com/abes-esr>
[3] <https://fr.wikipedia.org/wiki/SPARQL>
[4] <https://stp.abes.fr>