

Synthèse IST 2015

Françoise Thibault

Les infrastructures en information scientifique et technique

Au cours des années 2000, plusieurs réflexions et évolutions ont convergé qui ont donné une place croissante aux infrastructures en information scientifique et technique (IST). Les politiques publiques de recherche par leur quête d'indicateurs de résultats à court terme ont conféré aux publications scientifiques (revues essentiellement) une place centrale dans l'évaluation. Cette orientation a été favorisée par la généralisation du phénomène de numérisation des publications scientifiques qui s'est accompagnée de la production d'outils scientométriques de plus en plus performants et le développement de grandes plateformes internationales de diffusion des publications. Dans le même temps, la Commission européenne a lancé, dans le cadre du programme ESFRI (European Strategy Forum on Research Infrastructures), une réflexion sur les infrastructures numériques. En France, la publication, en 2008, du rapport de Jean Salençon essentiellement consacré à la circulation et l'accès aux savoirs scientifiques, la mission confiée à Bruno Racine par le Conseil du Livre pour « l'élaboration d'un schéma numérique pour les bibliothèques » et la mise en œuvre du processus d'élaboration d'une feuille de route sur les infrastructures de recherche (2008) ont conduit à inscrire l'IST comme un domaine à part entière doté d'infrastructures fondamentales pour la pratique scientifique.

Plusieurs objectifs étaient, d'emblée, visés : la volonté de sortir la documentation, l'information et la communication scientifiques des marges où elles ont été souvent placées, marges très problématiques pour l'économie globales des systèmes de recherche ; le souhait de structurer un secteur composé de membres souvent très actifs mais contraints dans leur action par l'éclatement du domaine ; l'ambition de limiter les coûts d'accès à une offre de ressources scientifiques éditées abondante, de qualité et répondant aux standards internationaux.

1- De la bibliothèque scientifique numérique (BSN) aux quatre plateformes HAL, OPEN-EDITION, COLLEX-PERSEE et NUMEDIF

Les deux premières feuilles de route ont retenu BSN comme infrastructure documentaire pivot. BSN vise deux objectifs principaux : répondre aux besoins des chercheurs en portant l'offre qui leur est fournie en IST à un niveau d'excellence mondiale et améliorer la visibilité de la recherche française.

BSN est composée de 9 segments d'activité pour lesquels des actions ont été initiées par différents opérateurs (spécialistes de la documentation, de l'IST, laboratoires de recherche...).

1. BSN 1 : L'acquisition des archives de l'édition scientifique et l'abonnement aux revues scientifiques (courant)
2. BSN 2 : Les dispositifs d'accès et d'hébergement aux/des ressources numériques (ENT type BIBLIOVIE, Biblio SHS, ...)
3. BSN 3 : Le signalement (bases de données bibliographiques)
4. BSN 4 : Open Access (construites notamment autour de la plateforme HAL)
5. BSN 5 : La numérisation du patrimoine scientifique de l'ESR national (revues et ouvrages) (construite autour de la plateforme PERSEE)
6. BSN 6 : L'archivage pérenne
7. BSN 7 : L'édition scientifique : coordination de l'aide publique, soutien à l'innovation de nouvelles formes d'édition et de diffusion et de nouveaux modèles économiques, etc.) (plateformes OpenEdition et NUMEDIF)
8. BSN 8 : La fourniture de documents/prêt entre bibliothèque
9. BSN9 : La formation et le développement des usages

Le travail, engagé en 2014, d'actualisation de la feuille de route des infrastructures de recherche a été l'occasion de dresser un bilan de la Bibliothèque Scientifique Numérique avec les acteurs de la Mission à l'information scientifique et technique, commune à la DGESIP et à la DGRI, en charge du suivi de ce secteur. Ainsi, depuis sa création BSN a effectivement permis de coordonner les différents acteurs dans les 9 domaines mentionnés précédemment. En terme d'actions collectives, BSN ne s'est pas investie de la même façon sur l'ensemble des secteurs et s'est centrée sur l'abonnement aux revues scientifiques, l'acquisition d'archives (projet « licences nationales ») et la relation avec les éditeurs scientifiques nationaux (projet « plateforme d'échange avec les éditeurs scientifiques»). Les licences nationales ont permis de dépasser l'émiettement institutionnel, elles ont pour objectifs de limiter les coûts des abonnements aux revues (actuellement + 6% par an) et de garantir l'équité entre chercheurs toutes structures de rattachement confondues. La plateforme d'échanges avec les éditeurs scientifiques a permis de réinstaurer un dialogue plus constructif avec les éditeurs scientifiques nationaux pour aider à dynamiser leur offre en matière d'édition scientifique.

Malgré ces résultats très positifs, BSN ne répond pas à l'ensemble des critères nécessaires pour prétendre au label « infrastructure de recherche ». La production de recommandations techniques ne suffit pas pour faire exister un opérateur et BSN ne peut faire état d'une gouvernance unique, d'un conseil scientifique et d'un budget global. Ces raisons ont conduit à identifier, dans BSN, les plateformes qui répondaient à l'ensemble de ces critères et de les inscrire en lieu et place de BSN dans la rubrique « Infrastructures en Information Scientifique et Technique ».

2- Les quatre plateformes HAL, OPEN-EDITION, NUMEDIF et COLLEX-PERSEE

Répondant à l'ensemble des critères des infrastructures de recherche, ces 4 plateformes constituent des ensembles complémentaires « usages-technologies » se succédant et interagissant dans le processus de diffusion de la production scientifique. Ces plateformes ont en commun la volonté d'ouverture à l'international de la production scientifique française et permettent d'assurer une certaine indépendance technologique tout en évitant le syndrome du « village gaulois ». Ces

plateformes interagissent toutes avec des plateformes internationales (métadonnées partagées, accords de partenariat, dispositifs de réversion).



2-1 La plateforme HAL : gestion des pré-prints, des multiples versions et archivage de la production scientifique en libre accès

L'archive ouverte HAL (<http://hal.archives-ouvertes.fr>) est une base de données en libre accès d'articles scientifiques et de données associées incluant TEL pour l'auto archivage des thèses. Initiée par le CNRS, cette plateforme a connu un développement important. En octobre 2006, les établissements de recherche et les universités ont signé un protocole d'accord en vue d'une approche coordonnée, au niveau national, pour l'archivage ouvert de la production scientifique. La plate-forme HAL est ainsi adoptée par l'ensemble des établissements d'enseignement et de recherche français (ou connectée aux plateformes spécifiques). Des règles d'articulation des systèmes d'archives locales institutionnelles avec la plateforme HAL ont été élaborées dans le cadre du dernier protocole d'accord. Ce document définit les modalités d'interopérabilité entre le système d'information de chaque établissement et la plateforme HAL. Les universités peuvent ainsi coordonner la mise en place d'outils de gestion des ressources numériques et plus globalement de leur système d'information avec la plateforme HAL.

Autour de cette plateforme, l'UMS (CNRS, Université de Lyon) réalise plusieurs autres outils : Sciencesconf.org - une plateforme Web qui s'adresse à tous les organisateurs de colloques, workshops ou réunions ; Episciences.org, plateforme technique d'examen par les pairs pour permettre l'émergence de revues électroniques en libre accès ; Héloïse, plateforme d'information des droits de diffusion des articles scientifiques de revues françaises. Ouverte aux collaborations internationales, la plateforme HAL est interconnectée à de nombreux réservoirs internationaux dont ARCHIV. Souple, elle permet aux universités comme aux organismes de recherche de bénéficier d'un portail personnalisé d'accès.

2-2 Les plateformes d'OpenEdition : diffusion en libre accès de la production scientifique édités et expérimentation de nouvelles formes d'écriture scientifique (revues et ouvrages essentiellement SHS avec outils associé).

Il s'agit aujourd'hui d'un ensemble de 4 plateformes regroupant 450 revues, 1000 carnets de recherche, 2000 livres renvoyant à plus de 28000 programmes scientifiques. OpenEdition est soutenu par le MESR depuis le début des années 2000 dans le cadre d'une politique de soutien à la diffusion scientifique en accès ouvert. OpenEdition est parvenu à fédérer les communautés de recherche au niveau international (près de 50 millions de visite annuelles). La mise en visibilité de cette réalisation

par son inscription dans la feuille de route nationale vise également à améliorer l'impact des projets de recherche pluridisciplinaires et leur capacité à transférer leurs résultats aux acteurs socio-économiques pour répondre aux défis sociétaux au niveau européen.

Il s'agit ainsi de :

- Améliorer l'accessibilité et la diffusion des publications scientifiques notamment en SHS ;
- Coordonner la transition numérique des acteurs stratégiques de la publication en SHS: institutions de recherche, bibliothèques, plateformes, éditeurs, financeurs
- Encourager les bonnes pratiques et établir des standards de qualité scientifique et éditoriale
- Mettre en place des mécanismes de financement durables, transparents, efficaces, pour la publication en libre accès
- Soutenir le développement de nouvelles méthodes de recherche par la publication numérique

2-3 NUMEDIF : diffusion et distribution multicanaux

La plateforme NUMEDIF est le résultat d'un travail de collaboration (issu de BSN) entre le service de diffusion de l'ensemble des presses universitaires (financés par le MESR depuis les années 1980 et confié à la FMSH) et une unité de recherche sur l'activité éditoriale de l'Université de Caen. Cette plateforme publique mutualisée de diffusion distribution multicanaux pour traiter des formes éditées multiples et adaptée à divers périmètres (volumétriques, géographiques ou de spécialité) repose sur la prescription de normes et standards dans le domaine de l'édition et de la diffusion numérique ; la définition de process et méthodes de constitution de contenus et de diffusion, actions de formation ; le développement d'outils de structuration de données éditées ; diffusion et formation à leur usage ; une animation de réseau de professionnels ; la construction de partenariats modèles en coédition et en diffusion distribution .

Cet appareil spécifique de diffusion-distribution vise à garantir la circulation des écrits scientifiques et leur mise à disposition optimale tant sous forme dématérialisée (archives, plateformes...) que matérielle (livres ou revues diffusés en librairie, auprès des bibliothèques ou des particuliers). Centré sur le livre en SHS, elle se doit d'atteindre les librairies de premier niveau et plus particulièrement les librairies à public universitaires pour la production courante, les bibliothèques universitaires et celles des centres de recherche et doit aussi être accessible dans le monde entier notamment au travers de la vente en ligne.

Une attention particulière est portée à la recherche d'articulations strictes avec les grands opérateurs publics du secteur, partenaires naturels : OpenEdition (plateformes), Hal (archives ouvertes), Persée (Rétro-numérisation)... tant en ce qui concerne la circulation des contenus structurés que celle des métadonnées... Le modèle et les outils d'informatique éditoriale qui l'accompagnent permettent déjà à partir d'une source unique de produire toutes les formes de diffusion. Il est d'ores et déjà compatible avec les schémas nécessaires pour les revues des plateformes Revues.org, Cairn et pour les livres diffusés sur Open Edition Books.

2-4 COLLEX- PERSEE : identification et consolidation de grandes collections scientifiques accessibles en ligne

La 4^{ème} infrastructure compète l'ensemble en faisant intervenir un monde trop souvent négligé du monde scientifique dans le processus de diffusion de la recherche à savoir celui des bibliothèques. *CollEx-PERSEE* rassemble dans un premier cercle des bibliothèques qui font référence dans leurs disciplines et s'inscrivent naturellement dans des stratégies à long terme : BDIC, BIU Santé, BNU de Strasbourg, BIU de la Sorbonne, GED Condorcet, BULAC, bibliothèques de l'INHA, de l'Observatoire de Paris, du MNHN, de la FMSH,...

Appuyée sur la plateforme de numérisation PERSEE, cette infrastructure permet de gérer et de valoriser des grands fonds scientifiques. Elle constitue un outil majeur des politiques de conservation partagée qui répondent à des logiques à la fois disciplinaires et de stratégies territoriales, en concertation avec la BnF.