



Direction de l'Information  
Scientifique et Technique

# **SCHEMA D'ORIENTATION STRATEGIQUE DE L'INFORMATION SCIENTIFIQUE ET TECHNIQUE (IST)**

## **"MIEUX PARTAGER LES CONNAISSANCES"**

**Constats, Vision, Principes d'Action**

**Novembre 2013**

Pour aller vers une Science publique ouverte à tous ses bénéficiaires, le CNRS s'engage à "Mieux partager les connaissances", ce qui conduit d'abord à mieux partager l'information scientifique et technique (IST). A l'heure numérique, assurer l'accès le plus large à l'IST est un défi aux dimensions paradoxales: les technologies ouvrent des voies prometteuses jugées irréversibles, et c'est l'option de la France comme celle de l'Europe; le quotidien du travail de la science nous montre toutefois que bon nombre de verrous doivent encore être levés à l'échelle globale. Comment y travailler?

L'IST ne fait qu'entrer dans une ère de changements multiples, qui concernent l'exploration la représentation, le partage des données et résultats de toutes les communautés scientifiques; les offres de procédés ou d'outils nouveaux se multiplient; les formes et les contenus de la documentation et de la publication se renouvellent, nécessitant parallèlement la mise à jour des règles, éthiques et juridiques, encadrant le travail de l'IST et de sa valorisation au bénéfice de tous les utilisateurs. L'accroissement des publications est universel et si le CNRS y joue un rôle moteur, c'est aussi en faisant écho aux préoccupations qui émergent chez les chercheurs de toutes les disciplines: comment se positionner dans la course poursuite engagée entre la croissance du volume des données et la pertinence de leur interprétation, comment offrir des services d'"Open Process" valorisant l'effort national et international d'Open-Access?

Ces questions-parmi bien d'autres, ont animé la démarche que le Collège de Direction du CNRS a voulu concrétiser dans ce Schéma d'Orientation stratégique, fruit d'une réflexion conduite avec tous les relais nationaux des acteurs de l'IST dans l'Organisme: pour eux tous, il est temps d'apporter des réponses construites aux défis que l'ère numérique introduit dans la construction de l'IST dont la science a besoin, aujourd'hui et demain. La réflexion stratégique initiée par la DIST (direction de l'information scientifique et technique) dans le cadre de ce Schéma d'Orientation, s'est donc donné les buts de préfigurer des réalisations immédiates comme des évolutions de moyen terme.

Prendre du recul n'avait de sens que pour déboucher sur les évolutions attendues par les chercheurs, et par tous les usagers de la recherche. Dans ce but, le Directeur général délégué pour la Science du CNRS lancera prochainement quatre Plans d'Actions Partagés (PAP) visant à mieux: 1."Se documenter", 2."Publier", 3."Analyser l'information", 4."Accompagner l'IST et sa valorisation". Les Instituts du CNRS relayeront cette démarche chacun selon leur choix, en poursuivant le travail engagé avec la DIST ([dist-direction@cns-dir.fr](mailto:dist-direction@cns-dir.fr)).

Ces quatre Plans d'Actions seront largement ouverts aux partenaires du CNRS pour que les Alliances, les Universités, les Organismes de recherche qui le souhaiteront, puissent concrétiser dans cet effort commun de partage de ressources et de projets, les liens forts qui les unissent au CNRS dans une démarche commune de chaque jour, pour bâtir une Science publique ouverte. Dans cet effort commun, où nous avons tous à gagner à "Mieux partager les connaissances," je souhaite que le CNRS soit une fois encore fidèle à la vocation nationale que lui confère son décret fondateur, qui est notamment de contribuer à "développer l'information scientifique".

Alain FUCHS,

Président du CNRS

L'Information Scientifique et Technique est décisive pour asseoir la compétitivité de la recherche française dans le contexte européen et international. En articulant toutes les disciplines et tous les partenaires de la recherche, l'architecture du CNRS est bien configurée pour tirer parti des avantages de la proximité et de la mise en réseau, que valorise l'IST à l'heure numérique : ces atouts spécifiques confortent la position de premier ensemble mondial par le nombre de publications scientifiques, que l'Organisme occupe à nouveau en 2013. Par ailleurs, la part relative du CNRS dans les publications scientifiques nationales et dans leur valorisation, continue de progresser.

L'IST englobe aussi bien les conditions dans lesquels les chercheurs ont accès aux données nécessaires pour mener leurs travaux, y compris lorsque celles-ci émanent directement du public, que les conditions dans lesquelles les résultats de leurs recherches sont mis à la disposition des communautés scientifiques et de toutes les attentes économiques et sociales. L'Information Scientifique et Technique est à ce titre un fondement solide pour la société de la connaissance que veulent construire la France et l'Europe.

De longue date, le CNRS a souligné le potentiel des outils numériques pour révolutionner la production de l'information scientifique et technique et pour favoriser le partage des résultats de la science<sup>1</sup>. Cette position est réaffirmée et approfondie par la Contribution du CNRS à la Stratégie Nationale de Recherche (juin 2013). Disposant d'une expertise reconnue, s'appuyant sur des personnels aux compétences unanimement admises, et engagé dans un partenariat fort avec les acteurs de l'enseignement supérieur et de la recherche, l'organisme est à même, à travers ce Schéma d'orientation stratégique, de tirer le meilleur parti du consensus qui se dessine mondialement pour un libre accès à tous des bénéfices de la recherche publique, pour un meilleur partage des connaissances.

« Mieux partager les connaissances », c'est d'abord mieux partager l'information scientifique et technique (IST), être à l'écoute des besoins présents des chercheurs et de la société.

Dans cette perspective, le Schéma d'orientation de l'IST du CNRS, résultat d'une réflexion stratégique interne,<sup>2</sup> a retenu un ensemble de pratiques, d'idées, de projets, débouchant sur:

**Un constat** : une dynamique de science publique ouverte

**Une vision** : une IST forte pour "Mieux partager les connaissances"

**10 Principes d'action** pour un "Plan d'Actions" à bâtir avec tous les partenaires du CNRS

---

<sup>1</sup> Le Décret du 24 Novembre 1982 portant sur l'organisation et le fonctionnement du CNRS précise (Article 2) que l'organisme a une vocation nationale à "développer l'information scientifique".

<sup>2</sup> La DGDS remercie tous les acteurs de la co-construction du Schéma d'Orientation stratégique de l'IST du CNRS, en particulier les 80 personnes environ mises à contribution dans la phase de correction, ainsi que le Comité de Lecture final, composé par: MM. Martino, Massiot, Sorger, Directeurs d'Instituts, Mme Michèle Leduc Présidente du Comité d'Ethique, Mme Marie-Claude Labastie, Secrétaire générale du Comité National, Mr Raymond Bérard, Directeur de l'INIST, Mme Christine Berthaud, Directrice du CCSD, M. Jean Daillant, Directeur TGIR Soleil, MM. Bruno David, François- Joseph Ruggiu, Ken Takeda et Mme Joanna Janik, Correspondants IST des Instituts.

## SOMMAIRE

Constat : Une dynamique vers la science publique ouverte .....	5
Pour une science publique ouverte .....	5
L'IST, vecteur d'une recherche publique ouverte .....	6
Défi 1: Ouvrir les conditions d'accès à l'IST .....	8
<i>Défi 2: Offrir une réponse cohérente à tous les besoins</i> .....	9
Vision : Une IST forte pour "Mieux partager les connaissances" .....	12
<i>Le CNRS est moteur dans cette démarche</i> .....	12
Le temps est venu de partager les ressources et les projets .....	14
Dix Principes pour l'Action d'IST .....	16
Principe d'action 1 : Anticiper et construire les options globales .....	17
Principe d'action 2 : Adapter les savoirs et les compétences .....	19
Principe d'action 3 : Former par et dans la recherche .....	23
Principe d'action 4 : Harmoniser l'offre et la demande d'IST .....	25
Principe d'action 5 : Partager les ressources communes HAL, Persée .....	32
Principe d'action 6 : Partager les données de la recherche .....	34
Principe d'action 7 : Identifier et Analyser les publications .....	37
Principe d'action 8 : Une stratégie globale pour le Libre Accès à la science publique .....	39
Principe d'action 9 : Partager l'éthique et les règles .....	42
Principe d'action 10 : Innover dans les démarches de l'IST .....	46

**Les pratiques et projets présentés par les acteurs de l'IST s'intègrent dans les principes d'action résumés ci-dessous.**

## **DIX PRINCIPES POUR L'ACTION IST**

### **UNE CONCEPTION EVOLUTIVE : PARTAGER LES SAVOIRS ET COMPETENCES AU SERVICE DE L'IST**

**PRINCIPE D'ACTION 1 : ANTICIPER ET CONSTRUIRE LES OPTIONS GLOBALES**

**PRINCIPE D'ACTION 2 : ADAPTER LES SAVOIRS ET LES COMPETENCES**

**PRINCIPE D'ACTION 3 : FORMER PAR ET DANS LA RECHERCHE**

### **DES MODALITES OUVERTES : ACCUEILLIR LES BESOINS DE TOUS LES CHERCHEURS**

**PRINCIPE D'ACTION 4 : HARMONISER L'OFFRE ET LA DEMANDE D'IST**

**PRINCIPE D'ACTION 5 : PARTAGER LES RESSOURCES COMMUNES HAL, PERSEE**

**PRINCIPE D'ACTION 6 : VALORISER ET PARTAGER LES DONNEES DE LA RECHERCHE**

**PRINCIPE D'ACTION 7 : IDENTIFIER ET ANALYSER LES PUBLICATIONS**

**PRINCIPE D'ACTION 8 : UNE STRATEGIE GLOBALE POUR LE LIBRE ACCES A LA SCIENCE PUBLIQUE**

### **UNE REGULATION PARTAGEE: DYNAMISER LES PRATIQUES**

**PRINCIPE D'ACTION 9 : PARTAGER LES REGLES**

**PRINCIPE D'ACTION 10 : INNOVER DANS LES DEMARCHES DE L'IST**

## CONSTAT : UNE DYNAMIQUE POUR LA SCIENCE PUBLIQUE OUVERTE

L'information scientifique et technique (IST) croît et se diversifie rapidement, dans un périmètre<sup>3</sup> qui comprend « l'ensemble des informations produites par la recherche et nécessaires à l'activité scientifique comme à l'industrie. De par sa nature, l'IST couvre tous les secteurs scientifiques et techniques et se présente sous de multiples formes : articles, revues et ouvrages scientifiques, spécifications techniques décrivant des processus de fabrication, documentation technique accompagnant les produits, notices de brevet, bases de données bibliographiques, littérature grise, banques de données brutes, archives ouvertes et entrepôts de données accessibles sur Internet, portails, etc. ».

Vecteur de la science ouverte, l'IST de l'ère numérique se confronte globalement à deux grands défis :

- Ouvrir les conditions d'accès,
- Offrir une réponse à tous les besoins.

Trouver des réponses à ces deux défis, qui servent de base à la démarche du schéma d'orientation stratégique de l'IST, c'est permettre de "Mieux partager les connaissances". Le CNRS entend être moteur dans ces évolutions, qui irriguent tous les projets IST de la recherche.

## POUR UNE SCIENCE PUBLIQUE OUVERTE

Plus que jamais, le CNRS est présent dans la construction d'une science publique ouverte, qui s'inscrit dans son projet fondateur : ses missions de recherche, de valorisation et de diffusion des résultats de la science sont en phase avec les enjeux de libre accès, enjeux liés en profondeur à ceux des mutations numériques.

---

## LE LIBRE-ACCES AUX RESULTATS DE LA RECHERCHE

La déclaration de Berlin<sup>4</sup> de 2003 sur l'Open Access, signée à l'époque par le CNRS, et portée par des institutions, définit le libre-accès *"comme une source universelle de la connaissance humaine et du patrimoine culturel ayant recueilli l'approbation de la communauté scientifique"*. Cette déclaration entend embrasser la totalité des productions de la recherche : *« Les contributions au libre accès se composent de résultats originaux de recherches scientifiques, de données brutes et de métadonnées, de documents sources, de représentations numériques de documents picturaux et graphiques, de documents scientifiques multimédia »*.

Au terme d'une décennie de négociations et de progrès des pratiques, l'année 2013 marque un net renforcement du consensus dans ce sens, sur des bases synthétisées en 2012 par le Comité d'Éthique du CNRS<sup>5</sup>.

---

<sup>3</sup> MESR, Service de la coordination stratégique et des territoires, Mission de l'Information scientifique et de la Recherche Documentaire (MISTRD)

<sup>4</sup> <http://openaccess.inist.fr/?Declaration-de-Berlin-sur-le-Libre>

<sup>5</sup> [www.cnrs.fr/comets/spip.php?article73](http://www.cnrs.fr/comets/spip.php?article73) , [www.cnrs.fr/comets/spip.php?article73](http://www.cnrs.fr/comets/spip.php?article73) , [http://www.cnrs.fr/comets/IMG/pdf/03-avis\\_relations-chercheurs-maisons-edition-2.pdf](http://www.cnrs.fr/comets/IMG/pdf/03-avis_relations-chercheurs-maisons-edition-2.pdf)

L'IST à l'heure numérique, c'est aussi le partage des données initiales et des données résultant de la recherche, accompagnées d'informations (les métadonnées) permettant de les réutiliser, à la fois pour valider les résultats et pour des objectifs différents des objectifs initiaux.

Dans sa déclaration du 24 janvier 2013, Geneviève Fioraso, ministre de l'enseignement supérieur et de la recherche, soulignait que, à l'heure numérique: " L'information scientifique est un bien commun qui doit être disponible pour tous". Dans cette perspective, on attend du libre accès à l'IST, qu'il permette l'amélioration de l'efficacité de la recherche, notamment en favorisant l'interdisciplinarité, la visibilité et l'impact des résultats pour la communauté scientifique elle-même mais aussi au-delà, pour les professionnels, les entrepreneurs, les journalistes et le public intéressé.

Les déclarations et prises de position se sont multipliées dans ce sens courant 2013: après Horizon 2020 et ses récents textes d'application, puis la déclaration des ministres chargés de l'ESR au G8 de Londres en juin 2013, de nouvelles prises de position se sont multipliées au Royaume Uni, aux Etats Unis, en Europe<sup>6</sup>.

---

## LE LIBRE-ACCES A L'HEURE NUMERIQUE

Dans la contribution du CNRS à la Stratégie nationale de Recherche (1<sup>er</sup> juin 2013)<sup>7</sup>, l'organisme se met en phase avec les enjeux et démarches de la révolution numérique. Le texte souligne notamment : "Aujourd'hui, la science vit une révolution (parmi d'autres) qui conduit à un nouveau paradigme selon lequel *la science est dans les données*. La production massive de données par des expériences scientifiques, des capteurs ou des populations munies d'équipement de communication (crowdsourcing) est dorénavant une approche classique". Par ailleurs, "L'extraction de connaissances, l'apprentissage, l'agrégation de données, la visualisation et la navigation dans de grands espaces de données sont autant d'instruments qui permettent d'observer et de découvrir des phénomènes, de valider des hypothèses et d'élaborer de nouveaux modèles".

## L'IST, VECTEUR D'UNE RECHERCHE PUBLIQUE OUVERTE

L'accès à l'information scientifique et technique, sa circulation et sa disponibilité sont des facteurs clés de l'efficacité de la recherche scientifique.

Comme aux premiers âges de la diffusion du livre, mais avec une vigueur incommensurable, l'IST numérique crée une nouvelle dynamique du partage des connaissances.

---

## L'IST A L'HEURE NUMERIQUE: UNE DYNAMIQUE NOUVELLE POUR LE PARTAGE DES CONNAISSANCES

L'Edition numérique facilite l'élargissement instantané du nombre des lecteurs, tout en permettant une réduction des coûts d'accès à l'IST ainsi qu'une souplesse d'exploitation et de distribution sans précédent. A coûts fixes souvent voisins de ceux de l'édition "papier", l'édition numérique présente l'avantage inédit de réduire très fortement-dans certains cas de supprimer-les coûts variables: la desserte d'un très grand nombre d'utilisateurs peut être envisagée sans risques d'exploitation ni dépense, tout en multipliant les services rendus aux usagers et l'enrichissement de contenu. Cette caractéristique

---

<sup>6</sup> Evolution de la position du Parlement britannique par rapport au rapport Finch, avec affichage d'une priorité au Green Open Access; publication des directives d'application de l'OSTP des Etats-Unis sur la mise à disposition publique des résultats de la science; Démarche Horizon 2020 en Europe; Déclaration de l'AAAS en faveur de l'adoption d'une refonte de la Charte des Nations Unies sur le partage des bénéfices de la science...

<sup>7</sup> [http://intranet.cnrs.fr/intranet/actus/img/SNR\\_Contribution\\_CNRS\\_final.pdf](http://intranet.cnrs.fr/intranet/actus/img/SNR_Contribution_CNRS_final.pdf)

technologique peut à elle seule expliquer la multiplication des formules de partage des connaissances développées ces dernières années à l'échelle globale, avec la création d' « Entrepôts documentaires universels ». <sup>8</sup>

---

### L'ERE DES "DEPOTS DOCUMENTAIRES ET DES CENTRES DE DONNEES UNIVERSELS"

En quelques années, la numérisation de la production a permis la mise en place d'un bon nombre de projets planétaires (Web of Science, par exemple, PLoS, Scopus, PubMedCentral ...) ainsi que des outils permettant de disposer d'une vue d'ensemble d'un sujet, de ses voies actuelles d'évolution, des découvertes intervenues dans les domaines connexes. A ces réservoirs d'information scientifique s'associent des métriques d'analyse dont la gamme s'étend constamment <sup>9</sup>: parmi d'autres, un inventaire global recense une centaine d'outils d'analyse répondant à toutes les grandes fonctionnalités associées à la démarche de recherche (recherche et comparaison d'idées, d'outils de calcul et d'analyse, d'outils d'exploration des préférences et opinions, d'outils d'exploration de corpus de données...). A travers ces outils se développent de nouvelles approches d'ingénierie des connaissances.

En parallèle, des centres de données disciplinaires, éventuellement mis en réseau entre eux et avec les publications, mettent leurs compétences thématiques au service de leur communauté et de l'ensemble des utilisateurs, scientifiques et autres : ils assurent la qualité des données, leur validité des données et leur capacité à être découvertes et réutilisées.

---

### LA POLITIQUE NATIONALE D'IST AU SERVICE DU CHERCHEUR : UN CADRE RICHE ET STRUCTURE

Au terme de réflexions et d'évaluations nombreuses auxquelles le CNRS a été constamment associé, la France s'est dotée d'une politique nationale d'IST numérique. Sous l'autorité du MESR, la Bibliothèque Scientifique Numérique (BSN), avec ses 10 Chantiers actuels<sup>10</sup>, permet d'accompagner les développements en cours de l'Édition scientifique en intégrant ses fonctionnalités complexes dans une approche globale. Le CNRS, qui est présent dans tous les Chantiers de la BSN à travers 24 experts, a fait récemment état de son analyse et de ses observations sur l'évolution en cours de la BSN: le Schéma d'orientation stratégique de l'IST du CNRS prend appui sur ses ressources et sur ses projets.

L'Agence Bibliographique de l'Enseignement supérieur (ABES), est un acheteur public et un expert de premier rang, partenaire de tous les usagers de l'ESR.

Le Consortium Couperin offre un cadre stratégique complet pour une démarche globale de négociation et de conseil dans la relation du monde scientifique (universités et organismes de recherche) avec ses partenaires et ses fournisseurs.

Le CCSD, en tant qu'opérateur national offre des solutions d'open access pour tous les chercheurs, enseignant-chercheurs via ses services centrés sur le dépôt et la diffusion en libre accès des publications dans HAL (Hyper Article en Ligne).

L'UMS PERSEE a pour objectif d'être un opérateur national pour la mise à disposition de la production scientifique rétrospective française, au service de la science en train de se faire (élargissant ainsi le programme PERSEE porté par le MESR qui était initialement principalement en direction des SHS, qui font elles-mêmes l'objet d'un investissement spécifique avec le TGIR Huma-Num).

Des initiatives pionnières ont conduit au développement de politiques d'IST cohérentes à l'échelle des organismes de recherche publics : regroupés dans l'Association EPRIST, les organismes de recherche mettent en commun leurs

---

<sup>8</sup> Celles-ci peuvent désormais relier les publications et les données avec le DOI (digital object identifier) dont l'INIST est l'Agence d'attribution en France.

<sup>9</sup> Voir par exemple: "Comparison of Research Networking Tools and Research Profiling Systems"  
[http://en.wikipedia.org/wiki/Comparison\\_of\\_Research\\_Networking\\_Tools\\_and\\_Research\\_Profiling\\_Systems](http://en.wikipedia.org/wiki/Comparison_of_Research_Networking_Tools_and_Research_Profiling_Systems)

<sup>10</sup> Le CNRS y est représenté par 24 experts, présents dans tous les Chantiers de la BSN.



expériences. Plusieurs grands établissements publics nationaux partagent ainsi les avancées des plus engagés d'entre eux<sup>11</sup>.

Le projet ISTE (Initiative d'excellence de l'Information Scientifique et Technique), dont le CNRS (DIST) est le pilote, vise à offrir à l'ensemble de la communauté de l'enseignement supérieur et de la recherche, "un accès en ligne aux collections rétrospectives de la littérature scientifique dans toutes les disciplines". Ce projet s'inscrit dans le programme "Investissements d'Avenir" initié par le MESR; l'ambition directe d'ISTEX est de "renforcer la recherche et l'enseignement supérieur français sur la scène mondiale"; ISTE permettra par ailleurs de valoriser la politique d'acquisitions de licences nationales, en rendant accessibles les collections courantes et les archives de façon transparente.

## DEFI 1: OUVRIR LES CONDITIONS D'ACCES A L'IST

A l'heure numérique, les modèles juridiques et économiques de l'IST sont soumis à rude épreuve: les règles de l'édition, du financement, de la diffusion de l'IST laissent place à l'incertitude. Le besoin global de clarification est manifeste. De cette clarification dépend la libre disposition des bénéfices de la recherche publique.

---

### LE DROIT DE L'EDITION ET DE LA DIFFUSION PUBLIQUE D'IST

Les usagers de l'IST au format numérique, les chercheurs et l'encadrement de la recherche, expriment un besoin de clarification des textes applicables à l'activité de publication et aux services qui y sont associés, notamment dans tous les cas où le droit public est frontalier du droit privé, mieux défini et défendu aujourd'hui.

Le droit de l'Édition scientifique publique, au regard de la liberté d'entreprendre, nécessite, comme en d'autres domaines du droit numérique, une évolution des règles et des usages.<sup>12</sup> Le Premier ministre impulse par ailleurs une action de régulation confiée au Conseil d'orientation de l'édition publique et de l'information administrative (COEPIA).

Les questions posées sont notamment celle du droit d'auteur, de la session des droits, des licences Creative Commons, etc

Le besoin de sécuriser la diffusion de l'IST dans la sphère publique grandit également au regard des multiples aspects du droit d'exploitation des archives numériques. Le CNRS est présent dans les évolutions de la diffusion publique des connaissances numériques, en particulier dans la conduite de l'Investissement d'Avenir ISTE: ce grand projet qui ambitionne de mettre à disposition des membres de l'ESR un ensemble inégalé de ressources numériques rétrospectives de la science, fait l'objet d'une évaluation juridique préalable, visant à conforter sa fonction d'infrastructure essentielle" au regard du droit public et du partage des connaissances au bénéfice de tous.

---

### LES REGLES ETHIQUES ET JURIDIQUES APPLICABLES AUX DONNEES DE LA RECHERCHE

Les "banques de données" sont partie intégrante de l'IST. Or la multiplication des bases de données, du partage et de la publication de données de la recherche, conduit à de nouveaux besoins, face auxquels les solutions sûres font parfois défaut. Un éventail croissant de communautés scientifiques s'interroge ainsi sur les conditions éthiques (respect de la vie privée, règles du partage, propriété intellectuelle...) et juridiques à réunir (responsabilité, déontologie, sécurité, droit de propriété, droit des libertés individuelles et collectives, droit commercial, droit pénal...), dans les opérations de stockage et de manipulation, de réutilisation et de partage des matériaux de l'IST que sont les données de la recherche (sciences de

---

<sup>11</sup> L'INRA en particulier, mais également l'INSERM, l'Inria, l'IRD, le CEA, l'Irstea...

<sup>12</sup> Dans ce sens, la BSN7 (Collège de l'Édition numérique scientifique) a diffusé en 2013, une "Charte des bonnes pratiques de l'édition scientifique numérique".

l'environnement, biologie, SHS notamment, mais également chimie, etc....). Ces clarifications<sup>13</sup> applicables aux séries de données sont des conditions pour faire face au besoin de partage des bénéfices de la science. ETALAB, en France, est aujourd'hui l'un des partenaires de valorisation et le destinataire potentiel d'une part de l'IST, en France. Cette politique nationale et européenne, fait l'objet de textes applicables en 2014.

---

## LES MODELES ECONOMIQUES DE L'EDITION

La marche vers l'Open Access s'est doublée d'une discussion générale sur le partage de la valeur dans la chaîne éditoriale, sur les marges associées à l'activité des groupes globaux, sur les modèles économiques optimaux de publication. L'hétérogénéité des formes de l'édition, tout comme la variété des rapports qu'entretiennent les communautés de publication avec l'Édition scientifique, proscrivent pour le moment le développement d'une solution unique d'application du principe de libre-accès à l'IST.

Avec tout un éventail de combinaisons intermédiaires, les trois familles d'approches présentes dans la diffusion de l'IST sont connues: celle des Revues dont le principe est le paiement par "l'auteur" (lequel est soit le producteur de la ressource, soit son laboratoire, soit son institution ou tout autre ayant-droit)<sup>14</sup>, soit le "lecteur" (ou tout autre usager d'une ressource en ligne achetée par une institution comme une bibliothèque par exemple). Enfin celles des revues en libre accès qui développent elles aussi plusieurs modèles économiques (modèle freemium d'Open Edition par exemple).

Une étude comparative des coûts de grande ampleur récemment publiée par Nature<sup>15</sup> distingue les publications associées aux principaux éditeurs scientifiques et celles relevant de l'Open Access, avec des coûts qui varient du simple à plus du double. Mais les différences entre les "modèles" doivent toutefois prendre en compte les services au chercheur présents dans chaque modèle; pour l'heure, les études manquent généralement sur le contenu en service des modèles, sur lesquels l'Académie des sciences mène actuellement une réflexion globale.

Cette réflexion préalable et nécessaire explique que la Commission Européenne ait pris position pour l'Open Access, sans trancher pour l'instant sur les modèles existants.

## DEFI 2: OFFRIR UNE REPOSE COHERENTE A TOUS LES BESOINS

Les besoins d'IST, différenciés d'une communauté scientifique à l'autre, sont -sauf exception-, en forte augmentation et promis à s'accroître. Des démarches d'« ingénierie des connaissances » se profilent et, avec elles, des voies et des modalités nouvelles de la publication scientifique.

---

## DES PRATIQUES EDITORIALES DIFFERENCIEES SELON LES COMMUNAUTES DE RECHERCHE

Tout en étant globalement acquises par principe à l'objectif de la science publique ouverte, les communautés scientifiques entretiennent des liens variables avec l'Édition scientifique : en biologie, par exemple, il est fréquent d'observer, simultanément, une démarche d'usage de grands dépôts d'Open Access comme PubMedCentral, tout en maintenant une

---

<sup>13</sup> En s'élargissant, la définition des données de la recherche peut comprendre aujourd'hui des ambiguïtés.

<sup>14</sup> D'où la discussion sur le montant des APC (Article-Processing Charges) que réclament les éditeurs pour contrepartie de leur coûts d'exploitation.

<sup>15</sup> Nature: "The true cost of science publishing. Cheap open access journals raise questions about the value publishers add for their money" Vol 495 28 March 2013, p.426

relation forte aux grands éditeurs internationaux voir remarque plus haut; les pratiques des communautés de la Physique accordent une part importante et grandissante aux grands dépôts de prépublications en Open Access, principalement arXiv, qui fait évoluer le positionnement de la revue par les pairs ; les astronomes entretiennent des liens étroits avec leurs publications les plus importantes, qui sont en petit nombre et gérées par des organes représentatifs de la communauté (sociétés savantes, ESO), tout en étant utilisateurs d'arXiv.

La communauté des mathématiques, grande utilisatrice d'arXiv et HAL, dispose par ailleurs d'un éventail de revues savantes souvent anciennes, encore sous forme papier, ce qu'elle a en commun avec bon nombre des communautés des SHS, par ailleurs très ouvertes et motrices dans les applications numériques de l'IST.

Certaines revues de ces deux communautés fonctionnent sur un modèle économique dont la viabilité est basée sur le concept de barrière mobile.

L'évolution prévisionnelle des options d'IST se doit donc d'être également présente à ces différentes voies d'accès à la publication, d'en saisir la dynamique systémique, d'en prévoir les conséquences de tous ordres, notamment budgétaires.

---

## LES PUBLICATIONS SCIENTIFIQUES : UNE AUGMENTATION CONTINUE ET FORTE

La croissance annuelle de la production scientifique globale est spectaculaire: en termes globaux, à partir des données Scimago (<http://www.scimagojr.com/>), la production annuelle globale de l'Édition scientifique est passée entre 1996 et 2012 de 1 134 000 à 2 250 000 articles/an, sur la base de Scopus, donc sans comptabiliser l'apport vigoureux des SHS. L'accélération est nette depuis 2004, avec l'arrivée en force des pays émergents dans le flux des publiants.

Une partie de cette augmentation est due à l'augmentation de l'activité scientifique au niveau mondial. Une autre proportion reflète cependant davantage l'augmentation de la pression à publier de plus en plus d'articles due au poids de plus en plus important accordé à la bibliométrie et amplifiée par les intérêts des éditeurs commerciaux. Cette composante de l'augmentation n'est donc pas forcément un signe de bonne santé de la science pour l'INSMI.

---

## DONNEES ET PUBLICATIONS : UNE COURSE POURSUITE

L'activité scientifique du 21<sup>ème</sup> siècle se construit sur une nouvelle dimension d'investigation, un nouveau paradigme : avec le développement des moyens d'investigation (capteurs, sondes, séquenceurs, instruments d'observation), les masses de données générées sont importantes et au-delà du défi de l'analyse et du traitement de ces données, la question de leur organisation, stockage et accès devient une préoccupation majeure.

Malgré sa croissance rapide, l'effort de publication ne croît pas en proportion des données produites par la recherche. En prenant pour base l'effectif global des deux millions de publiants mondiaux actuels, un ordre de grandeur<sup>16</sup> a été donné sur la destination finale des données d'observation produites par la recherche : 2% des données prennent la direction de dépôts numériques spécialisés (PetDB, PDB, SedDB, MiRB, TAIR...); 8% des données scientifiques sont acheminées vers de grandes bases généralistes (DataVerse, Dryad, dépôts institutionnels). Ainsi, 10% des données de la recherche seraient utilisées par les publications scientifiques et 90% des données resteraient stockées sur les disques durs locaux. Certaines disciplines pionnières ont mis en place une politique de données "ouvertes" : en astronomie par exemple, les données observées par les télescopes sol et spatiaux sont mises à disposition de tous après une période "propriétaire" (en général un an), sous la responsabilité des agences en charge des instruments.

---

<sup>16</sup> Une évaluation récente, parmi d'autres, offre une image du volume total des observations produites par le travail scientifique annuel global. A. de Waard: "Epistemic Modality and Knowledge Attribution: Types and features" Elsevier Labs, Universiteit Utrecht, July 2012

## VERS UNE "INGENIERIE DES CONNAISSANCES"

Avec l'essor des coopérations internationales, de la spécialisation à l'interdisciplinarité, du financement par projets, du recours massif aux grands équipements et aux moyens de calculs partagés, les pratiques de recherche ont considérablement évolué ces dernières décennies. Il y a une dizaine d'années a été forgé le concept d'escience (enhanced science, John Taylor, 1999), faisant référence à cette nouvelle science collaborative et distribuée, s'appuyant sur l'accès à des grands réservoirs de données, des réseaux à haut débit et des moyens de calcul et de visualisation à haute performance. L'ensemble des acteurs, financeurs, gouvernements, organismes de recherche, sociétés savantes, éditeurs, opérateurs, industriels sont appelés, comme l'analyse le récent rapport de la Royal Society (2012) à concourir au développement d'une « Science as an open entreprise ». Ce rapport met l'accent sur la nécessité de s'attaquer au déluge des données par des technologies modernes afin de réaliser, au bénéfice de la science et de la société, leur extraordinaire potentiel.

L'efficacité de la science repose aujourd'hui sur sa capacité à traiter et exploiter les masses de données produites dans des environnements partagés; son efficience dans les années à venir se mesurera à sa capacité à gérer, partager et autoriser la réutilisation de ces mêmes données.

Les "banques de données brutes", partie intégrante de l'IST dans la définition ministérielle, sont ainsi utilisées dans des démarches, inégalement intégrées entre elles, qui conditionnent ensemble l'activité et l'architecture de publication.

Une mise en cohérence de ces choix n'apparaît pas relever de l'incongru dans le Schéma directeur de l'IST du CNRS, pour aller vers une démarche lisible d'"ingénierie des connaissances" qui s'esquisse actuellement. Les questions génériques qui se posent ainsi de façon transverse aux disciplines portent sur les normes et standards de description des données (métadonnées), les règles de partage, d'accès sécurisé, de stockage et d'archivage pérenne, en addition aux aspects juridiques et d'éthique abordés plus haut.

Il y a ainsi autant de sujets stimulants pour cette "ingénierie des connaissances" dont les outils, par ailleurs, se développent rapidement dans une gamme très étendue<sup>17</sup> à base de fouille de texte et de données, d'ontologies, de traitements du web sémantique. L'objectif pourrait être de fournir au chercheur comme aux utilisateurs, un service intégré d'appui à la construction des publications numériques.

Les personnels de soutien à la recherche du CNRS, au plus près des chercheurs dans les unités, ou dans les structures de service, participent déjà au développement et au maintien des infrastructures de données, tout en développant des approches qui visent à réunir les fonctionnalités développées ci-dessus (INSHS, INSU notamment).

---

<sup>17</sup> Voir par exemple: "Comparison of Research Networking Tools and Research Profiling Systems"

## VISION : UNE IST FORTE POUR "MIEUX PARTAGER LES CONNAISSANCES"

Avec des résultats probants au regard de son organisation, le CNRS confirme son action d'entraînement en matière d'IST à l'heure numérique. Cette action peut prendre appui sur une politique nationale d'IST, qui fournit un cadre d'amont favorable. Le temps est donc venu de partager les ressources et les projets avec les équipes de recherche des universités et organismes: le Schéma d'orientation stratégique fournit un cadre pour bâtir ensemble le Plan d'Actions pour l'IST.

## LE CNRS EST MOTEUR DANS CETTE DEMARCHE

### DES RESULTATS DE PUBLICATION PROBANTS A L'HEURE NUMERIQUE

Comme cela est connu, l'architecture de l'organisme comporte une double articulation, celle de toutes les disciplines et celle de tous les partenaires de la recherche: le CNRS bénéficie ainsi à la fois d'une proximité de ressources et de projets dans presque tous les grands domaines de la recherche et peut ainsi optimiser les effets de réseau que requiert la Science ouverte. Cette double articulation s'associe à des résultats- moins connus- qui démentent l'analyse selon laquelle une forte recherche fondamentale s'opposerait au partage amont-aval : le CNRS a déposé plus de 400 brevets en 2012 et se place ainsi au 5<sup>e</sup> rang national<sup>18</sup>.

A l'heure numérique, ces atouts de double articulation peuvent servir à interpréter le positionnement favorable du CNRS dans le champ de la publication scientifique globale.

Comme en 2012, le CNRS se situe une fois encore dans les classements de 2013, au premier rang des institutions scientifiques mondiales en matière de documents de recherche publiés et répertoriés dans les bases de données internationales (Scimago Institutions Rankings, SIR), et ceci avec un taux de co-publications internationales de 58%.

Ce positionnement global est par ailleurs dynamique: dans le classement 2013 du SIR, les critères de qualité des publications associés au rang du CNRS sont majoritairement en progrès. D'autres classements mondiaux spécialisés corroborent cette première image; par exemple, le Nature Publishing Index 2012 place le CNRS au 5<sup>e</sup> rang mondial des institutions de recherche dont les résultats sont publiés par les revues du groupe Nature, cette position est en hausse (7<sup>e</sup> rang en 2011).

### UN ROLE D'ENTRAINEMENT DANS LE PARTAGE ET LA VALORISATION DES CONNAISSANCES

Le CNRS progresse par ailleurs nettement dans la part relative des publications issues d'unités liées au CNRS.

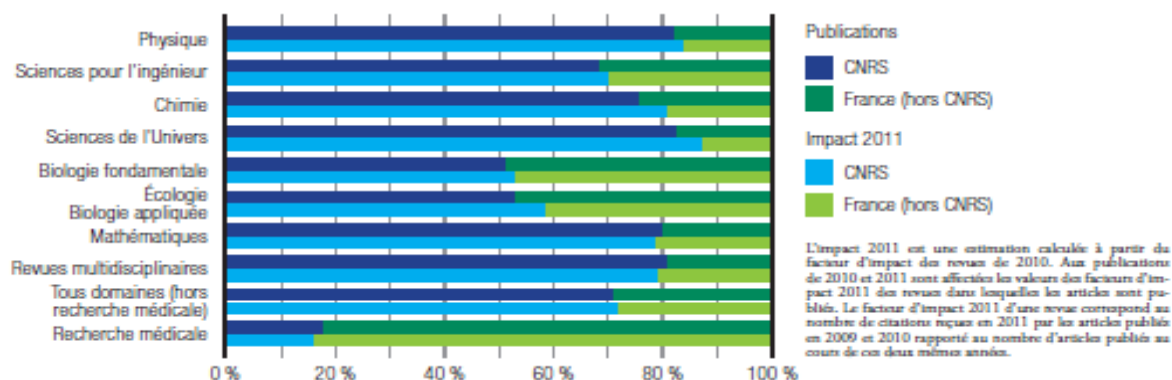
L'effet moteur du CNRS, à l'amont des publications scientifiques nationales, est prépondérant et dynamique. Le diagramme ci-dessous en donne un aperçu : la part des laboratoires liés au CNRS est toujours au moins égale à 50%, souvent supérieure à 70% selon les secteurs, toujours croissante ces dernières années, ce qui procure une mesure globale de "l'efficacité" des moyens mobilisés par l'organisme.

---

<sup>18</sup> Voir le palmarès des principaux déposants de brevets publiés en 2012 - Chiffres clés 2012 de l'INPI [http://www.inpi.fr/fr/presse/espace-presse/communiques-de-presse/detail-communique/article/5068.html?tx\\_ttnews\[backPid\]=1983&cHash=fb836f72c1&S=&print=1](http://www.inpi.fr/fr/presse/espace-presse/communiques-de-presse/detail-communique/article/5068.html?tx_ttnews[backPid]=1983&cHash=fb836f72c1&S=&print=1)

## Contribution des laboratoires liés au CNRS à la production française de publications scientifiques

Part des publications des laboratoires liés au CNRS et leur impact dans la production scientifique française par domaine des sciences de la matière et de la vie (2010-2011)



Source: Données Thomson Reuters. Traitement CNRS / SAPPS, 2013

Toutes disciplines confondues (hors recherche médicale), la part des publications des laboratoires liés au CNRS dans la production scientifique française atteint 70,9% en 2011<sup>19</sup>. Ce résultat est constamment en progrès ces dernières années. Par les collaborations qu'elles attestent, ces données placent l'organisme en situation motrice dans le développement des actions de soutien à l'IST, sur lesquelles le CNRS a été souvent précurseur.

### UN ROLE PRECURSEUR EN MATIERE D'IST

L'action du CNRS en la matière est ancienne et diversifiée, même si elle a pu manquer de visibilité. L'évaluation du CNRS par l'AERES notait en effet en 2012<sup>20</sup> : "Un effet de fédération des initiatives et des énergies paraît attendu de la participation du CNRS au projet national de bibliothèque scientifique numérique (BSN) mais presque tout reste à faire pour que le CNRS utilise mieux l'immensité des données et des connaissances disponibles en son sein." Pourtant, c'est bien au CNRS, comme déjà indiqué, qu'ont pris place les premières positions institutionnelles en faveur de l'Open Access. D'autres réalisations majeures, parfois bien antérieures, sont également à mentionner comme le Centre de données astronomiques de Strasbourg (CDS-INSU) qui a fêté en 2012 ses 40 ans d'existence, la Chimiothèque nationale (INC), le soutien aux bases de données par la création de l'unité BBEES à l'INEE, le Livre blanc Calcul intensif, et plus récemment l'initiative Mastodons portée par la Mission Interdisciplinarité du CNRS (MI), avec notamment le projet PREDON.

L'InSHS de son côté, dispose des TGIR Huma-Num et Progedo pour porter une réflexion structurante à l'échelle de tout l'Institut, c'est-à-dire sans restriction concernant les sous disciplines. L'enjeu d'Huma-Num est bien de bâtir une infrastructure pour les humanités numériques, allant même au-delà de la gestion des données pour prendre en compte également les moyens d'analyse et de traitement de ces données, Progedo assurant pour sa part la gestion des données quantitatives. Par ailleurs, le CLEO est une unité très innovante en matière d'édition électronique en libre accès avec plus 350 revues en Sciences humaines et sociales, 700 carnets de recherche et une plateforme dédiée aux collections de livres inaugurée en février 2013.

<sup>19</sup> Contribution du CNRS à la stratégie nationale de Recherche, 1<sup>er</sup> juin 2013, p.6: "Le rôle stratégique des SHS dans l'interconnexion des défis sociétaux permet d'éviter la segmentation et la redondance, d'assurer un maillage par la recherche de base des questions sociétales pour lesquelles une mobilisation interdisciplinaire est nécessaire(...)"

<sup>20</sup> AERES Evaluation du CNRS, 2012, Note 61

A partir d'une enquête sur les bases de données existantes au sein des laboratoires en 2010, l'INEE a pris conscience de l'urgence à intervenir pour éviter la perte d'une partie d'entre elles ; d'où la création de l'UMS BBEES qui vise à apporter soutien et méthode dans la gestion des bases de données de recherche sur la biodiversité.

Chaque Institut peut ainsi citer une sous-discipline qui a mené une réflexion sur le sujet des données, car certaines communautés, qui souvent partagent des équipements lourds producteurs de données, suivent des standards et se sont défini des pratiques communes (astrophysiciens, sismologues, géochimistes, biologistes moléculaires,...). Pour leur part, les mathématiciens se préoccupent de la disponibilité de leurs productions anciennes, pertinentes dans cette discipline, à travers NUMDAM.

---

### DES RISQUES A MAITRISER

Face à ses atouts, le CNRS a pour défi de faire face, fortement, aux besoins de renouvellement des connaissances et qualifications de ses agents : les Réseaux des personnels IST sont prêts à agir dans ce sens et se sont préparés à le faire. Il y a, en outre, une marge de progression importante à réaliser dans une meilleure articulation entre les personnels IST et les chercheurs, dans les communautés où cette relation n'a pas encore pu se mettre en phase, ainsi que sur la sensibilisation des chercheurs à la démarche du partage de leurs données.

Parmi les risques également, existe celui de l'accumulation incontrôlée de nouveaux dispositifs, là où ceux-ci seraient déployés sans mise en commun et sans évaluation : le développement, très rapide, des nouvelles démarches, nécessite à l'évidence, une mise en cohérence et un partage.

---

### LE TEMPS EST VENU DE PARTAGER LES RESSOURCES ET LES PROJETS

---

#### UNE REFLEXION STRATEGIQUE INTERNE

Pour situer les enjeux de l'IST, comme l'y invite sa mission nationale en la matière<sup>21</sup>, le CNRS a donc entrepris de dresser l'état des savoirs, des analyses et des ressources présents dans l'organisme, souvent depuis longtemps. La DGDS a donc confié la mission à la direction de l'information scientifique et technique (DIST), de produire un état des lieux puis d'accompagner la mise en place d'une stratégie du CNRS, définie en relation avec toutes les communautés de travail de l'IST.

Au terme de cette démarche, le Collège de Direction du CNRS a retenu les options du présent Schéma d'Orientation stratégique, résultat d'une réflexion co-construite<sup>22</sup> avec les communautés de l'IST dans l'organisme<sup>23</sup> autour d'environ 80 experts et décideurs appartenant aux communautés scientifiques du CNRS.

---

#### UN PLAN D' ACTIONS A CO-CONSTRUIRE AVEC LES UNIVERSITES ET LES ORGANISMES

Dès décembre 2013, sous l'impulsion de la DGDS, cette réflexion stratégique interne pourra déboucher sur un Plan d'Actions IST, que le CNRS propose à ses partenaires (Universités, Organismes de recherche, Alliances, OST)<sup>24</sup>, de co-

---

<sup>21</sup> Le Décret du 24 Novembre 1982 portant sur l'organisation et le fonctionnement du CNRS précise (Article 2) que l'organisme a pour vocation, parmi ses missions nationales, à "développer l'information scientifique".

<sup>22</sup> La DIST remercie les contributeurs directs de cette démarche et les membres du Comité de rédaction du document final (.....)

<sup>23</sup> Les dix Instituts du CNRS ainsi que plusieurs de leurs composantes IST internes (Cléo, MathDoc...), les réseaux de personnels de l'IST (RBDD, Medici, Renatis), les Unités spécialisées d'IST (INIST, CCSD, Persée), le Comité d'Ethique et le Secrétariat général du Comité national. La DIST a également recueilli les contributions de deux des TGIR (Huma-Num et Soleil). Deux directions du siège (DASTR et DSI) et la Mission pour l'interdisciplinarité ont par ailleurs apporté leur appui, ainsi que le Bureau de Bruxelles du CNRS (DGRI). Des contacts ont par ailleurs été établis dans ce cadre, avec ETALAB, le GENCI et le COEPIA.

<sup>24</sup> Le CNRS compte 845 UMR sur 1053 unités de recherche.

construire en faisant jouer tous les effets de levier des partenariats qui sont construits, pour mieux partager l'IST. Ce Plan sera également décliné en fonction des calendriers contractuels du CNRS et de ses partenaires, (Contrat avec le MESR 2014-2017, Conventions de Sites). Ce Plan d'Actions s'adossera en amont aux ressources et travaux en cours des Chantiers de la Bibliothèque Scientifique Numérique (BSN), et contribuera ainsi à en diffuser les acquis dans les actions opérationnelles.

Le CNRS soumet donc les dix principes d'action qui suivent, à la discussion et à la préparation de choix communs avec ses partenaires.



## DIX PRINCIPES POUR L'ACTION D'IST

### TROIS CLEFS POUR L'ACTION D'IST

Dix principes pour l'action d'IST se dégagent de la réflexion menée avec les acteurs de l'IST que nous avons articulés autour de trois grandes clefs.

#### UNE CONCEPTION EVOLUTIVE

L'IST évolue très vite dans ses problématiques et dans ses besoins : la prospective y tient une place essentielle pour situer les besoins dans le changement rapide des réseaux des dispositifs et des outils. Par ailleurs, la pratique de l'IST est d'abord affaire de qualification et de compétences des chercheurs et des personnels techniques qui en ont la responsabilité : la formation et l'échange des pratiques tiennent donc une place essentielle. La conception de l'IST doit pouvoir évoluer et s'adapter.

#### DES MODALITES OUVERTES

Les actions d'IST doivent faire face aux besoins de tous les chercheurs et de toutes les communautés de bénéficiaires: cette contrainte implique de formuler des modalités ouvertes. L'approche des modalités qui a été ici retenue est celle de la filière de la production d'IST: ainsi sont successivement passées en revue l'offre d'IST du CNRS telle qu'elle est adossée aux ressources présentes dans les Chantiers de la BSN, dans les unités spécialisées de la DIST (INIST, Persée, CCSD), mais également l'offre présente dans les Instituts, ou encore dans les grandes ressources externes (ressources de calcul, grandes bases partagées de type ISTE), qui sont autant de conditions d'accès à l'offre d'IST. Les modalités de l'IST, ce sont également les choix d'accès à la publication des diverses communautés de la recherche et les soutiens qui leur sont apportés. Ce sont, enfin, les dispositions assurant la traçabilité et le répertoire des publications scientifiques. La démarche de Libre-accès est présente à tous ces niveaux tout en faisant l'objet d'une stratégie globale de moyen terme. Cette démarche soulignant l'INSMI et l'INSU, doit se faire avec la participation des chercheurs et des solutions innovantes doivent être explorées.

#### UNE REGULATION PARTAGEE

La régulation de l'IST repose sur l'appropriation des règles de production et de partage, et en premier lieu des règles éthiques et juridiques. La régulation intervient ensuite dans le partage des métriques d'analyse, pour tous les besoins de suivi des Instituts à des fins d'évaluation et de pilotage. La régulation dépend, par ailleurs des avancées que procurent les outils innovants, les réseaux d'innovation en IST. La régulation d'ensemble des projets IST fait également appel, en fonction des besoins, à des choix de gouvernance (celui d'un Comité de régulation de l'IST à la DIST a été évoqué pour effectuer le suivi du Plan d'Actions IST).

## UNE CONCEPTION EVOLUTIVE : PARTAGER LES SAVOIRS ET COMPETENCES AU SERVICE DE L'IST

### PRINCIPE D'ACTION 1 : ANTICIPER ET CONSTRUIRE LES OPTIONS GLOBALES

#### L'ACCUMULATION DES BESOINS

Il n'y a pas que les données qui fassent "déluge" : les projets associés à l'IST se développent avec une vigueur souvent torrentielle.

- L'accumulation des besoins se traduit par une demande globale de repères sur la direction, l'organisation et la portée des démarches d'IST, et en particulier sur le positionnement à retenir dans les évolutions de moyen terme.

Ainsi, la **TGIR SOLEIL**, qui accueille et fait cohabiter bon nombre des grandes communautés des autres disciplines, retient le besoin d'un positionnement épistémologique à construire sur l'IST produite à base des données numériques et des nouveaux outils d'analyse et de partage des connaissances.

L'**INEE** l'observe, il y a une surabondance d'informations : il faut bien y distinguer les flux "entrants" et les flux "sortants" d'IST, pour lesquels les problèmes se posent de façon distincte. Quelle est la bonne solution d'accès à l'IST aujourd'hui et demain ? Plusieurs sources d'IST ou une source unique regroupant les autres ? Il n'y a pas de cohérence d'ensemble et il faut la rechercher : le CNRS a la capacité d'y contribuer, ce qui nécessite de prendre du recul.

L'**INC**, suivie en cela par une majorité d'Instituts, se montre favorable à une approche analytique et stratégique des publications des Instituts, par la mise au point d'outils appropriés de métrique. L'**INEE** a mené une première expérimentation dans ce sens.

L'**IN2P3** et l'**INP** préconisent ensemble, à travers des initiatives comme SCOAP3, le besoin d'aller vers le libre-accès par une démarche globale construite à partir d'expérimentations immédiates, solides et partagées.

A l'**INS2I**, on voit se dessiner l'idée qu'il faudrait une fonction d'appui à l'ensemble de la recherche d'outils et de modélisation des sources de données et de publications.

L'**INSIS** souligne la nécessité d'explorer de nouveaux domaines scientifiques à l'échelle du monde, de développer une veille et de la bibliométrie pour entrer en profondeur dans la maîtrise des outils de la valorisation, des brevets notamment, et de remettre à plat pour cela l'analyse des dispositifs nationaux existants.

L'**INSB** fait état d'une relation construite à la fois sur le recours aux archives ouvertes (PubMedCentral), sur la publication systématique dans les grandes revues internationales, sur l'accès à de grandes bases de données : de nombreuses questions éthiques et juridiques posées dans les sciences de la vie mériteraient une approche construite au niveau du CNRS.

L'**INSHS** souligne le besoin de repères partagés, de règles communes, dosant l'incitation et l'obligation, fondant le partage de l'IST sur une vision claire des modèles disponibles et accessibles à court et à moyen terme sur les grands thèmes stratégiques de l'IST. Une feuille de route est nécessaire, pour les chercheurs comme pour les professionnels de l'IST, par-delà les affirmations de principe, de façon à faire progresser les grandes dossiers comme le libre-accès, l'accès à l'information scientifique et la valorisation de la production scientifique.

L'**INSMI**, qui recommande la construction d'une approche numérique intégrée et globale de l'IST (de l'édition à la recherche de données), constate qu'il y a nécessité de faire progresser et de partager les connaissances sur les structures de données, sur l'hétérogénéité des observations, condition de base pour le développement des informations produites par l'IST. Pour autant, il est nécessaire de faire cohabiter dans la durée (et donc de financer) les bibliothèques papier et les spécialistes de ces documents avec les nouvelles bibliothèques numériques sous leurs diverses formes

Pour l'**INSU**, il faut construire des « piliers solides » pour faire vivre la transversalité et l'adhésion des disciplines. Il organise par exemple ses services de données des Sciences de la Planète en "pôles de données" thématiques, en collaboration avec les autres organismes impliqués (CNES, Ifremer, etc), et a nommé un chargé de mission pour suivre la problématique des données à l'échelle de l'Institut.

---

#### UN PROJET PROSPECTIF COHERENT

- Les positions du CNRS sont motrices dans la définition et la promotion des positions nationales, européennes et internationales en faveur d'une IST partagée.

Comme les constats globaux l'ont mis en évidence, le CNRS est présent dans toutes les enceintes nationales, européennes et internationales où progresse le projet de la Science ouverte. Ces positions peuvent s'appuyer sur des ressources et des réflexions internes qui développent autant de lignes d'actions opérationnelles. Les difficultés à assurer la cohérence globale des positions en faveur de la science ouverte ne sont pas sous-estimées, comme le montre les étapes en cours des négociations sur l'Open Access (voir principe d'action 8).

- Une réflexion interne se dégage en faveur d'une mise en œuvre opérationnelle de ces options.

Un effort global de communication scientifique est à construire. Le **Cléo**<sup>25</sup> souligne la nécessité de structurer une vision politique nationale et européenne sur plusieurs bases, qui conditionnent ensemble la modernisation du système de communication scientifique.

Le Cléo observe également qu'il convient ainsi d'agir en introduisant "des innovations radicales dans l'écosystème" : promotion d'un Libre-accès reposant sur de véritables incitations envers les chercheurs (évaluation, dotations, valorisation), avec un principe de bonus pour l'accès ouvert, promotion de l'« open peer-review » dans la mesure du possible, promotion également de nouvelles formes éditoriales (carnets de recherche, par exemple), modernisation des dispositifs technologiques sous-tendant le système de communication scientifique dans l'identification des documents (DOI) et des auteurs (ORCID, Idref, ISNI...), en définissant des dispositifs centrés sur l'utilisateur, sur l'interopérabilité entre les instruments (en s'inspirant des acquis d'Huma-Num) et sur la pertinence des indicateurs d'IST, sur l'archivage pérenne.

La **TGIR Huma-Num** met en œuvre une démarche globale pour les SHS : la TGIR porte en effet l'ERIC (European Research Infrastructure Consortium) de Dariah (Digital Research Infrastructure for the Arts and Humanities).

L'**INIST** dispose d'une riche expérience des produits et services de l'ingénierie documentaire pour le soutien et l'accompagnement, le développement d'outils documentaires, les normes et standards afin de favoriser l'interopérabilité des plateformes et services. La **DIST** envisage début 2014, en liaison avec l'INIST et les Instituts du CNRS, la mise en place d'une mission de contacts aux Etats-Unis, sur le thème des ressources et projets en matière de prospective de l'IST (prévisions, évaluations, travaux en cours).

- Les positions internationales, européennes et nationales en faveur de la Science ouverte sont confortées par l'ampleur des besoins et la qualité des propositions de mise en œuvre. Ces positions s'appuient à l'amont sur l'action du MESR dans le cadre de la BSN.

Le Schéma d'Orientation stratégique s'appuie sur le projet national de la BSN, dont le programme observe : "Le secteur de l'édition scientifique a été totalement bouleversé par les nouveaux modes de diffusion du savoir, à la fois sur les plans économique et heuristique.(...) Aux médiations traditionnelles a succédé un ensemble d'usages qui méritent d'être encouragés tout en conservant la liberté indispensable à l'exercice scientifique et la qualité des protocoles de validation de la production scientifique". Le Chantier BSN 7 mène dans ce sens une veille globale sur l'Edition scientifique numérique et a mis en place la ressource nationale qu'est le Collège de l'Edition numérique scientifique. Une enquête nationale est lancée sur l'Edition scientifique : le CNRS y participera activement.

---

<sup>25</sup> Centre pour l'Edition Electronique Ouverte <http://cleo.openedition.org/>

## 1.1 CREER UNE PROSPECTIVE DE CHOIX IST DE MOYEN TERME

En s'appuyant sur sa réflexion interne, en la partageant avec ses partenaires, le CNRS peut envisager d'étoffer une démarche prospective sur l'IST portant, sur les stratégies d'achat et d'accès aux ressources, sur les savoirs partagés en bibliométrie (BSN 3), sur la politique nationale de l'édition scientifique (BSN7), sur la prise en compte de la transversalité de l'open access (BSN4). Cette démarche s'appuiera sur sa réflexion interne, l'expertise de ses opérateurs, sur les acquis précieux des organismes de recherche partagés dans le cadre de l'association EPRIST, et sollicitera le concours de l'ABES et du Consortium Couperin. Au terme de sa construction, cette démarche devra permettre de connaître et de situer les évaluations et travaux internationaux et nationaux relatifs aux perspectives à moyen terme de l'IST notamment sur:

- l'évolution des politiques documentaires des grands organismes de recherche et universités à l'étranger,
- les outils d'analyse des corpus de publications et données,
- les partenariats de production des services d'IST à la recherche et à sa valorisation,
- les normes internationales applicables aux documents, données, bases...
- les stratégies internationales et nationales des grands organismes de recherche étrangers en matière d'IST.

Cette liste est non limitative. Les thèmes sur lesquels portera cette démarche seront définis en relation avec les Instituts, afin de permettre au CNRS de disposer rapidement des résultats de cet investissement indispensable au développement de sa vision à moyen terme des projets IST à construire et à partager avec ses partenaires.

## PRINCIPE D'ACTION 2 : ADAPTER LES SAVOIRS ET LES COMPETENCES

Adapter les compétences et les savoirs est évidemment essentiel dans un univers où les changements de l'IST ont été, sont, et seront vraisemblablement encore, très rapides. La question se pose aux niveaux des personnels d'IST comme à celui des chercheurs, qui éprouvent également un besoin d'appui méthodologique.

### L'ACTION DE FORMATION ET D'ECHANGE DE SAVOIRS

Le CNRS bénéficie du concours d'un personnel formé aux actions d'IST, et relevant de plusieurs familles professionnelles, dont la révision est en cours au MESR avec le concours actif du CNRS (OMES).

L'Organisme bénéficie par ailleurs des synergies qu'occasionne la mise en réseau très dynamique de ces personnels, notamment dans les réseaux nationaux que sont : Renatis, Medici, RBDD, RNBM, Démocrite et Docplanets. Ces réseaux ainsi que les divers réseaux régionaux et thématiques des personnels, sont particulièrement présents dans la réflexion nationale sur le devenir de l'IST.

Cette réflexion, il est vrai, rejoint pour eux la préoccupation professionnelle d'évolutions de carrière et de qualification qu'impose l'évolution des modèles économiques et technologiques de l'IST. Dans ce contexte, les Réseaux ont pu apporter au CNRS, à la fois, l'expérience très riche et diversifiée de personnels présents dans toutes les communautés de la recherche, mais aussi, de façon particulièrement opportune et bénéfique, une expertise à jour sur les nouveaux modèles de l'IST, notamment ceux associés aux données de la recherche.

Dans ce sens, les personnels d'IST se sont montrés d'emblée parties prenantes des mutations en cours et prennent une part active à la définition de besoins de formation dans lesquels s'intègre spontanément la perspective et l'environnement de travail des laboratoires et la relation aux besoins des chercheurs : ce sont là des opportunités dont il faut évidemment souligner l'intérêt dans le cadre de ce Schéma d'Orientation stratégique, auquel les Réseaux ont activement participé. Ses suites doivent être approfondies en relation étroite avec les DU.

## 2.1 REEVALUER LES QUALIFICATIONS ET LES PROFILS D'EMPLOIS

### Constats

Au 31 décembre 2012, les personnels d'IST travaillant au CNRS présentent un large éventail de compétences, avec les caractéristiques globales suivantes :

- un personnel conséquent (770 ITA CNRS et 132 non CNRS) clairement identifié IST (familles A et C de la BAP F<sup>26</sup>)<sup>27</sup>,
- des personnels n'étant pas identifiés IST mais assurant des fonctions informatiques indispensables à l'IST (certains personnels de la BAP E notamment) ou travaillant spécifiquement en IST pour un institut scientifique très spécialisé (certains personnels des familles A, B et C de la BAP D<sup>28</sup> par exemple),
- des personnels de niveau élevé de qualification (la moitié environ des personnels CNRS de la famille A de la BAP F est IE-IR) et ayant souvent une double voire une triple compétence (scientifique, informatique, IST) ; des profils très variés : certains IR, (mais aussi AI et IE, selon l'INSU), effectuent la saisie des publications, selon l'INEE ; les correspondants IST en SHS sont le plus souvent des ITA BAP F mais il y a également des BAP D et des chercheurs.

Ce personnel, composé en majorité de femmes, est inégalement réparti au sein des Instituts comme entre les unités. La moitié environ des personnels engagés dans les activités d'IST a plus de 50 ans, avec localement des disparitions progressives de la fonction (dans les UMR INSU notamment). Des exemples pionniers, comme celui du CDS, montrent l'impact important de documentalistes de la BAP F formés à la prise en charge des données scientifiques dans la mise à disposition des résultats de la recherche.

Pour InSHS l'éventail d'analyse des qualifications s'étale sur les 3 BAP D, F et E. Il y a là un chantier stratégique RH à bâtir autour de la formation, de l'évolution des métiers et du développement des carrières, les trois notions étant liées. La question du devenir des personnels de moindre qualification est une question sensible et importante.

### Vision

L'enquête de la DIST rejoint l'opinion établie selon laquelle il y a localement une difficulté à mettre en synergie les compétences des professionnels de l'IST et les besoins des chercheurs : l'idée a été souvent exprimée que ces deux catégories d'acteurs de l'IST avaient le besoin d'approfondir, voire de renouer, une relation de collaboration efficace, à l'initiative des DU. Les personnels expriment nettement le souhait de faire évoluer leurs carrières et leurs qualifications au service de l'IST.

### Principes d'action 2.1

Le groupe BSN9 intitulé "Formation, Compétences et Usages" répond à la nécessité de prendre en compte les questions de formation liées aux différents chantiers de la BSN. Il a travaillé jusqu'ici sur le signalement de l'offre publique de formation à l'Information Scientifique Numérique et à l'élaboration de fiches de compétences de l'information scientifique, dans le cadre du référentiel national des Emplois-types de la Recherche et de l'Enseignement Supérieur (REFERENS) :

<sup>26</sup> BAP F : Documentation, édition, communication. Famille A : IST, collections patrimoniales. Famille C : édition multi supports, impression, graphisme [http://www.dgdr.cnrs.fr/drh/omes/documents/pdf/RapportIST\\_2011.pdf](http://www.dgdr.cnrs.fr/drh/omes/documents/pdf/RapportIST_2011.pdf)

<sup>27</sup> Source : Labintel au 31/12/2012, traitement CNRS / SAPPS

<sup>28</sup> BAP D : SHS. Famille A : traitement et analyse de bases de données. Famille B : traitement, analyse et représentation de l'information spatiale. Famille C : analyse de sources écrites, iconographies et orales <http://www.dgdr.cnrs.fr/drh/omes/documents/pdf/Etude-BAP-D-2012.pdf>

- Redéfinir les métiers des professionnels de l'IST en relation avec BSN9 et REFERENS (sans oublier les personnels de moindre qualification) et partager l'offre de formation, notamment universitaire,
- Mettre en place des actions renouvelant la relation avec les chercheurs en mettant en avant la valeur ajoutée apportée par les professionnels de l'IST.

---

## 2.2 DEFINIR UN PLAN NATIONAL DE FORMATION POUR L'IST

Vues les caractéristiques des personnels du CNRS, il n'y a pas d'alternative à une remobilisation au service de l'IST : c'est de toute évidence un investissement nécessaire et profitable. Les Réseaux des personnels IST ont répondu à la DIST dans le sens d'un Plan national de formation IST, qui pourrait être l'occasion de partenariats multiples au service de l'IST. A l'InSHS, le réseau des correspondants IST dans les laboratoires avec aujourd'hui 183 représentants désignés par le directeur du laboratoire pour 280 laboratoires (liste de diffusion, site web <http://corist-shs.cnrs.fr>), est une force essentielle pour démultiplier l'action.

### Constats

Des réseaux métier très actifs existent :

- au niveau national (Renatis<sup>29</sup>, Medici<sup>30</sup>, RBDD<sup>31</sup>, URFIST<sup>32</sup>),
- au niveau régional (ARPIST, GO!Doc, Isidora, Mistral, Doccitanist, ...),
- thématique (DEMOCRITE, RNBM, Docplanets...).

Ces Réseaux se concentrent sur les problèmes relevant des évolutions du métier, de la prospective en IST et de l'adaptation des outils, des compétences et des services proposés aux chercheurs. Ces aspects sont notamment traités à travers les actions de formations, comme les FRéDoc<sup>33</sup>, les ANF ou les groupes de travail. Les offres de formation sont nombreuses et diversifiées sur le plan interne au CNRS (INSHS, opérateurs IST) comme externe (URFIST notamment). Les Instituts ont une pratique de formation qui peut être partagée, comme celle de l'InSHS, qui met en place 4 à 5 ANF par an.

Les réseaux régionaux et le réseau national Renatis se concentrent « sur les actions conduisant aux évolutions du métier, de la prospective en IST et de l'adaptation des outils, compétences et services proposés aux chercheurs ».

Le Réseau Medici propose de conduire une cartographie nationale sur les métiers et compétences de l'édition scientifique publique, avec établissement d'un plan de formation pluriannuel. Ce réseau propose également une offre de service aux DU (aide à la rédaction des fiches de postes, aide au recrutement temporaire, suivi des agents entrants, propositions de formation pour stimuler la carrière des agents de la BAP F, famille C).

### Vision

Les réseaux de professionnels de l'IST sont une force et un atout pour diffuser les bonnes pratiques, fédérer et former les communautés. Ils sont actifs et proches du terrain donc susceptibles de faire remonter les besoins des chercheurs. Ils sont partie prenante dans l'évolution des métiers de l'IST et les besoins en formation des professionnels de l'IST. Une vigoureuse campagne de formation aux langues étrangères en général et à l'anglais en particulier, est de mise pour parvenir à une

---

<sup>29</sup> Réseau national de l'information scientifique et technique (IST) du CNRS <http://renatis.cnrs.fr/>

<sup>30</sup> Réseau national en édition scientifique publique <http://medici.in2p3.fr/>

<sup>31</sup> Réseau données et bases de données <http://rddb.cnrs.fr/spip.php?rubrique1>

<sup>32</sup> URFIST : Unité Régionale de Formation à l'Information Scientifique et Technique : <http://urfist.enc.sorbonne.fr/>

<sup>33</sup> FRéDoc 2013, par exemple, sur le thème de "Gestion et valorisation des données de la recherche".  
<http://renatis.cnrs.fr/spip.php?rubrique78>

coopération quotidienne entre les acteurs européens de l'IST. Un soutien et une formation juridique paraissent également importants.

### Principes d'action 2.2

- Soutenir les réseaux métier notamment pour réussir l'articulation avec les communautés de terrain, autour d'un pilotage par les DU.
- Définir un plan de formation national en s'appuyant sur ces réseaux, en relation étroite avec les correspondants IST des Instituts (corist), BSN9 et les acteurs de la formation en IST en interne au CNRS avec des opérateurs tels que l'INIST, Huma-Num, ... ou en externe (l'URFIST, le CNAM et certaines Universités qui proposent des formations spécialisées).

## DES ACTIONS SPECIFIQUES

---

### 2.3 LANCER UNE ACTION NATIONALE SUR LES DONNEES DE LA RECHERCHE

En matière de gestion des données de la recherche, les besoins de développement sont immenses et nécessitent de mobiliser le trio professionnels de l'IST, informaticiens et scientifiques.

Certains instituts ont mis en place des organisations visibles et performantes, par exemple :

- l'INSU a créé le CDS<sup>34</sup> il y a 40 ans pour assurer la collecte et la dissémination des données astronomiques et informations associées pour la communauté internationale et dispose de documentalistes en son sein (AI, IE et IR),
- l'INEE a mis en place BBEEES<sup>35</sup> qui conseille ou intervient dans les unités, quelques jours ou quelques mois pour relancer, restructurer les bases de données Biodiversité, Ecologie, Environnement et Sociétés,
- Pour l'InSHS, la TGIR Huma-Num<sup>36</sup> favorise par l'intermédiaire de consortiums regroupant des acteurs des communautés scientifiques, la coordination de la production raisonnée et collective de corpus de sources (recommandations scientifiques, bonnes pratiques technologiques).
- A l'INSB, ces formations sont déjà bien cadrées, en liaison avec les informaticiens (bioinformatique).

L'utilité des professionnels de l'IST est avérée dans d'autres domaines que la gestion des données de recherche tels que la mise en œuvre du Web des données par exemple (présentation des travaux de l'INRA aux FRéDoc2013). L'INS2I suggère la mise en place d'une structure de data management plan avec diverses étapes, ce qui permettrait de posséder non pas simplement des modes d'emploi mais aussi des normes communes en termes de curation.

### Vision

Les besoins sont très variables au sein des laboratoires et les solutions doivent être adaptées en conséquence, comme l'observe l'INEE. Certaines communautés scientifiques sont bien organisées ou s'organisent en matière de gestion de données et leur expérience peut être d'une grande utilité.

---

<sup>34</sup> Centre de Documentation de Strasbourg <http://cdsweb.u-strasbg.fr/index-fr.gml>

<sup>35</sup> L'unité Bases de Données sur la Biodiversité, Ecologie, Environnement et Sociétés <http://bbees.mnhn.fr/> a pour objectif de structurer et d'optimiser le travail autour des bases de données de recherche sur la Biodiversité naturelle et culturelle, actuelle et passé.

<sup>36</sup> Huma-Num <http://www.huma-num.fr/>

Le CNRS a des atouts majeurs pour bien gérer les données de recherche en termes de compétences nécessaires (IST, informatique, scientifique). Des moyens sont à trouver pour pallier l'inégale répartition des professionnels de l'IST sur le territoire. (Équipes mobiles? Formations complémentaires? Centres régionaux de ressources?)

### Principes d'action 2.3

- S'appuyer sur les pratiques et le savoir-faire des équipes déjà constituées et multiplier les expériences concrètes en matière de gestion des données de la recherche, dans un pilotage appuyé sur les DU,
- Définir et mettre en place les organisations et moyens nécessaires pour pallier l'impossibilité de généraliser des équipes multi-compétences (IST, informatique, scientifique), physiquement en un même lieu (équipe volante, plateforme collaborative, centralisation des données dans des centres spécialisés...)
- Mieux appréhender les besoins des communautés (INEE a émis l'idée de lancer des appels d'offre nationaux pour faire ressortir les besoins en matière de gestion des données de la recherche par exemple).

---

## 2.4 UNE PRIORITE FORTE POUR LA FORMATION A L'INIST

L'INIST est en train de concevoir le Plan de développement des qualifications qui accompagnera le redéploiement d'une partie de ses activités vers de nouveaux projets. L'INIST a enfin le besoin de faire face aux adaptations nécessaires des savoirs acquis dans ses tâches historiques. Un projet formation de grande envergure est donc à prévoir.

En parallèle l'INIST est apporteur de formation vers les personnels des unités.

L'INIST vise le soutien des chercheurs et professionnels pour acquérir de nouvelles compétences et optimiser leur savoir-faire en IST par le biais de formation en présentiel et à distance ou par le développement de supports de type e-learning sur une plateforme partagée, mais également le développement d'une ingénierie de la formation.

- Le séminaire INIST a par ailleurs confirmé le besoin important de formations, veille et alertes portant sur l'ensemble du spectre de la production, gestion et traitement des productions scientifiques, données de la recherche incluses. L'étendue de ce spectre nécessite l'implication de compétences réparties dans l'Institut. L'organisation et l'animation d'un réseau de formateurs en interne sont donc nécessaires pour répondre efficacement aux demandes.
- L'INIST serait également en capacité d'intervenir dans les formations initiales ou continues vers les communautés (école doctorale et chercheurs) et vers les professionnels de l'IST.

## PRINCIPE D'ACTION 3 : FORMER PAR ET DANS LA RECHERCHE

Il y a généralement peu de formations à l'IST pour les chercheurs, et celles-ci, qui relèvent de la formation par la recherche, sont inégalement réparties. Ces formations sont présentes dans des écoles doctorales et des laboratoires dont la cartographie reste à réaliser.

Par ailleurs certains des services offerts par les opérateurs IST du CNRS sont mal connus des chercheurs.

L'InSHS, par exemple, estime qu'il serait utile de sensibiliser les chercheurs sur les grands enjeux des problématiques de l'IST, les former aux nouveaux outils IST, les conseiller pour faire les bons choix sur les technologies de l'information, les accompagner au niveau juridique (bonnes pratiques, modèles de contrats, etc.). L'INEE avance l'idée d'appels d'offres nationaux pour répondre aux besoins et offrir des formations ciblées.

Dans cette catégorie d'offre nationale, Huma-Num avance l'idée de nouveaux services et d'actions pilotes, reposant sur des relais de formation pour entrer de plain-pied dans le web des données, le *linked data*. Le chantier est essentiel. "Le travail de structuration des acteurs nationaux offrant des formations sera décisif (Huma-Num, Cléo, INIST, CCSD). Les initiatives



développant des guides de bonnes pratiques (Huma-Num, Cléo, BSN) doivent être non seulement encouragées mais également reprises à d'autres échelles" souligne le Cléo.

Une meilleure appréhension par les chercheurs des enjeux et des moyens à mettre en œuvre pour le partage des données vise à la fois à faire évoluer les pratiques individuelles, mais aussi à préparer une évolution des pratiques d'évaluation de l'activité scientifique, qui doivent accompagner l'évolution du paysage si l'on veut obtenir une participation active de la communauté : il faut arriver à faire reconnaître l'impact du partage des données, et le temps passé à préparer celles-ci, comme partie intégrante du travail scientifique, au même titre que la publication d'articles dans les revues à lecteurs, par tous les chercheurs, y compris ceux qui participent aux instances d'évaluation.

---

### 3.1 STRUCTURER LA DEMARCHE DES ACTEURS NATIONAUX

- Mettre en place des actions de sensibilisation et formation en IST auprès des doctorants (écoles doctorales) et des chercheurs au sein des laboratoires.
- Structurer et articuler le travail des acteurs nationaux offrant des formations (Huma-Num, Cléo, INIST, CCSD, Persée, URFIST, OST ...) et développer des initiatives de conseils en bonnes pratiques.

## DES MODALITES OUVERTES : ACCUEILLIR LES BESOINS DE TOUS LES CHERCHEURS

A tous les niveaux de la production des publications scientifiques, l'action d'IST vise à apporter une solution aux besoins de chaque communauté scientifique, tout en créant la dynamique de pratiques communes améliorant la lisibilité, la traçabilité, la visibilité, le partage des résultats de la science. Ces démarches sont toutes parcourues par l'objectif transversal du libre-accès aux bénéfices de la recherche publique, pour lequel par ailleurs est développée une stratégie globale à moyen terme.

### PRINCIPE D'ACTION 4 : HARMONISER L'OFFRE ET LA DEMANDE D'IST

Harmoniser l'offre et les besoins d'IST suppose un double mouvement d'adaptation entre les Instituts et l'INIST, qui n'ont eu qu'inégalement à s'accorder dans leurs projets. A l'heure où l'INIST est en phase avec une politique nationale de l'IST qui attend beaucoup de son offre d'IST, la demande des Instituts est le repère de réorientation de ses choix, comme l'a voulu son récent séminaire de travail. La relation évolue également entre les autres porteurs des projets d'IST, au premier rang desquels figure le CCSD et, tout récemment PERSEE. Comme l'INIST, ces deux derniers acteurs de l'offre d'IST au CNRS, visent à être en phase avec tous les besoins de l'enseignement supérieur et de la recherche.

### LA DEMANDE DES INSTITUTS: L'EVOLUTION DES BESOINS

L'accès à l'information, l'économie de la publication, la valorisation de l'IST sont trois thèmes sur lesquels les Instituts ont concentré leurs réponses. Celles-ci, autant que possible sont à classer en retenant la distinction émise par l'INEE, entre flux entrant et flux sortant d'IST.

#### 4.1 FACILITER L'ACCES AUX SOURCES DOCUMENTAIRES

Faut-il un portail unique? La question de ce premier "flux entrant" qu'est l'information reçue *pour* la recherche, est généralement abordée avec un souci de simplification. Bon nombre de chercheurs utilisent peu d'outils et surtout celui dont ils ont l'habitude pour le travail quotidien. Les utilisateurs du Web of Science (WoS) forment un ensemble construit et acquis aux logiques d'accès à l'information par portail global.

Pour l'INC, face à la nette augmentation de la production scientifique, le numérique permet la facilité de transmission, l'instantanéité mais aussi la multiplication des formats possibles (publications dans les revues, production de données, logiciels, blogs, ...). L'INS2I, l'INEE, l'IN2P3, l'INSB, l'INSIS pratiquent les revues électroniques et n'ont pas d'appétence pour le papier. Mais l'INC souligne également que les chercheurs ont accès à un nombre parfois trop élevé de portails : le temps d'accès de recherche et de gestion des portails est une contrainte supplémentaire dans l'accès à l'information. L'INSB, qui est utilisateur du Portail "BiblioVie" de l'INIST, fait état des redondances de cet outil avec les portails des universités et surtout celui de l'INSERM.

Une large majorité des Instituts a exprimé le souhait d'un accès UNIQUE, rapide, ergonomique, et pratique, à l'IST : il faut "rendre l'outil invisible". L'influence des grands moteurs globaux est centrale (le WoS Google, bien sûr, sont constamment cités). L'organisation actuelle proposée par l'INIST, est trop "tubulaire" en ce qu'elle n'autorise pas d'accès croisé aux portails. Une minorité d'Instituts, tel l'IN2P3 ou l'INSB, déclarent n'avoir pour l'instant pas engagé de relation suivie avec l'INIST.

La communauté mathématique travaille sur la mise en place d'un portail par le biais de la cellule MathDoc, du réseau RNBM et du réseau Mathrice (informaticiens des unités de mathématiques). Ce portail, conçu pour être un "portail à façon", spécifique à la discipline, s'appuiera sur le hub de métadonnées des collections acquises ou utiles. Il n'est pas conçu comme un concurrent au portail unique mais bien comme un complément, s'appuyant sur des bases communes et offrant des services spécifiques à la discipline plus large que les services documentaires (notamment des services informatiques et collaboratifs proposés par Mathrice). Ce projet s'appuie sur l'expertise de la cellule MathDoc qui a joué un rôle moteur dans un projet européen similaire (EuDML).

#### Principes d'action 4.1

- Mettre en place une architecture d'accès à l'IST répondant aux évolutions des besoins d'accès à l'information des chercheurs des Instituts et de toute la recherche.
- Développer préalablement l'analyse de répartition des flux de choix et de dépenses (Instituts, Portails, Universités...) à partir du chantier lancé par la DIST et les Corist avec la DDAI et la DSFIM.

---

#### 4.2 L'Economie de la publication

##### A. OPTIMISER ET PERENNISER L'ACCES AUX PUBLICATIONS PATRIMOINE SCIENTIFIQUE

Comme l'observe l'INC, "le site de l'éditeur est la bibliothèque du chercheur". Pour l'INEE, "le WoS est l'archive des chercheurs".

- La relation avec les grands éditeurs et leurs intermédiaires doit s'inscrire systématiquement dans les démarches nationales portées par la DGRI, avec le concours stratégique de l'ABES et du Consortium Couperin.
- Faire évoluer et systématiser les négociations d'établissement et la stratégie des achats du CNRS, en relation avec la MISTRD, avec l'appui de l'ABES et du Consortium Couperin. Faire évoluer le modèle d'affaire pour optimiser les services aux chercheurs (simplification d'accès, clarification de l'offre des portails, des services d'analyse).

##### B. CONCEVOIR UNE DEMARCHE STRATEGIQUE DE PRESENCE SUR LES ARCHIVES OUVERTES

arXiv est très utilisé par les mathématiciens mais il y a une nécessité que le CNRS ait une relation forte et construite avec arXiv. Ainsi, il existe, pour l'INSMI, un risque d'une relation de dépendance mal contrôlée : la publication scientifique ne peut être placée dans les mains d'une entité indépendante comme ArXiv, hébergée à l'université Cornell. Le problème est d'ordre stratégique car cette base est unique. Les miroirs qui existaient en Europe sont en train de disparaître. Si l'accès à ArXiv rencontre des difficultés, le patrimoine scientifique peut se trouver menacé. Pour l'INSU et l'INSMI le maintien de la relation à arXiv est important : une convention paraît nécessaire. Il faudrait une négociation globale et siéger au CA. Heureusement, le CCSD gère un des rares miroirs pour les dépôts français d'arXiv.

- Consolider les possibles partenariats avec les grandes Archives ouvertes privées et publiques (ArXiv, ...) et les réseaux sociaux (Research Gate...)
- Développer un régime juridique protecteur de l'archive ouverte, inspiré du droit public (infrastructures essentielles ?)

##### C. PRESERVER LE PATRIMOINE SCIENTIFIQUE ANCIEN

L'archive de la science est loin d'être réductible au seul domaine numérique et les collections doivent être préservées dans toutes les directions où l'IST est présente. Plusieurs disciplines (mathématiques, physique notamment, mais aussi SHS), signalent l'importance de l'archive ancienne comme source nécessaire au développement de la science en cours de construction.

Un savoir-faire existe en matière de numérisation au sein de l'INSMI (Mathdoc via le programme NUMDAM) : il faut le maintenir car il est reconnu y compris au niveau européen. C'est en effet Mathdoc qui a imposé le standard de métadonnées définies pour le projet EUDML. Cela correspond aussi à un besoin car des ressources fondamentales en mathématiques ne sont pas encore numérisées et ne sont pas couvertes par le projet ISTEEX.

A l'InSHS, le programme Persée du MESR qui était consacré à la numérisation d'archives scientifiques en sciences humaines et sociales est élargi avec le portage du CNRS, de l'ENS et du CUE de Lyon, sous l'égide de BSN5. L'INSU a également des besoins, notamment avec les revues géologiques. Cette catégorie de besoin ne concerne pas que les publications : l'INSB, par exemple, participe aux bio-banques européennes de stockage et des bases de données de séquençage.

A l'IN2P3, à l'InSHS, et à l'INP, mais aussi à l'INSMI et le RNBM, le papier reste un support recherché : il est important de conserver un voire deux exemplaires papiers pour la France entière, ne serait-ce qu'en attendant que des solutions d'archivage numérique stables et fiables soient mises en place. Le RNBM travaille (en partenariat avec le CTLe) à la mise en place d'un plan national de conservation partagée des périodiques papiers pour assurer une pérennité de ces collections, et permettre leur circulation via la fourniture de documents. Il s'agit de limiter les doublons mais aussi d'organiser au plan national le maintien systématique et la conservation pérenne de plusieurs collections des titres les plus importants dans la discipline.

- Construire un programme national cohérent et interdisciplinaire de préservation du patrimoine scientifique ancien incluant un archivage pérenne (en liaison avec BSN6).

#### D. ASSURER LA PERENNITE DES EDITEURS ACADEMIQUES DE REVUES DE QUALITE

Le passage à l'Open Access est rendu difficile pour les revues et éditeurs académiques : bon nombre de sociétés savantes se refusent à faire payer l'auteur.

L'INC et l'INSMI sont favorables à la formulation d'une Charte des relations avec les éditeurs des sociétés savantes, fondée sur les notions d'excellence, de complémentarité (besoin non satisfait par d'autres sphères éditoriales), de poids du CNRS dans les achats et dans la publication. A l'INSMI, la production éditoriale académique nationale et internationale est riche et abondante et nécessite une prise en compte.

- Définir une position commune, après enquête sur la taille des besoins. Participer activement à l'enquête nationale de la BSN7.
- Se référer aux recommandations de l'INC et à la Charte des bonnes pratiques développée par le Chantier BSN7.
- Soutenir les projets émergents (Episciences.org (math et informatique) le Journal de l'Ecole Polytechnique...) et/ou subventionner les revues, mutualiser les savoir-faire.

#### E. CONFORTER LA NEGOCIATION DES ACQUISITIONS LOCALES

L'InSHS choisit de négocier lui-même ses revues avec CAIRN, pour rester ferme sur le périmètre des revues à acquérir. La suggestion est faite d'inclure au besoin un chercheur dans le périmètre des négociations. L'InSHS entend définir ses acquisitions et abonnements au plus près des besoins des chercheurs et des laboratoires. Le

RNBM a tenté de le faire, en se maintenant dans la négociation avec Springer en y amenant un regard scientifique. Il y a un besoin de soutien juridique aux instituts qui s'engagent dans les négociations

L'INSMI est en train de monter une organisation de mutualisation afin d'assurer la pérennité des services offerts par les 3 sociétés savantes (françaises et européenne) les plus proches de la communauté.

Au-delà des observations critiques sur les "bouquets" proposés par les éditeurs, la demande générale des Instituts est de se rapprocher davantage de la demande du chercheur, et ceci par une démarche construite déclinable par les DU.

- Mettre en place, avec le concours de la DSFIM, des outils et procédures pour assurer la visibilité et la simplification des démarches locales d'acquisition documentaire.
- Partager ces démarches avec les partenaires du CNRS dans les laboratoires.

#### 4.3 VALORISER L'INFORMATION SCIENTIFIQUE

Sujet très vaste, la valorisation est abordée ici sous l'angle de questions immédiates : archives ouvertes, services et outils créés pour le chercheur inégalement mobilisés dans le quotidien de la recherche. Une démarche nationale claire est souvent recommandée. La question centrale de l'affiliation des publications, fait l'objet d'un développement séparé.

##### A. VALORISATION PAR L'INSTITUTION

L'IN2P3 exprime, tout comme les autres Instituts, une option de principe favorable au dépôt dans HAL, (l'INSIS et l'INSB expriment des réserves). L'InSHS soutient fermement le dépôt dans HAL. La démarche toutefois suscite des interrogations et suggestions précises. Outre les questions de pérennité évoquées plus haut, le dépôt dans les archives ouvertes soulève deux ordres de questions : l'opportunité d'une démarche nationale, la qualité des services associés au dépôt. Ainsi, l'INEE, qui alimente HAL, observe que toute la production scientifique n'y est pas recensée, que le service rendu se limite actuellement aux recensements des publications précédemment déposées. L'INP ainsi que l'INSU font observer la nécessité d'améliorer des "vases communicants" simples entre HAL et arXiv, afin d'assurer une représentation pertinente et complète des dépôts et de valoriser ainsi le projet national, soutenu par le MESR et l'Académie des sciences.

Face à cette démarche nationale d'accès à la science ouverte, la valorisation des publications passe par plusieurs recommandations concordantes :

- Mettre en place une réflexion de l'organisme sur les règles de dépôt dans HAL : doit-on aller vers un dépôt automatique, systématique? Selon quelles règles applicables localement dans les laboratoires?
- Quels services d'appui à l'Open Access développer dans une optique de services d'appui au dépôt dans les archives ouvertes ("Open Process")?

##### B. VALORISATION PAR DES OUTILS CREEES POUR LE CHERCHEUR

Le chercheur se voit proposer de nombreux outils de valorisation de leur production nés sur le marché privé dans les communautés du libre, dans les réseaux sociaux (Research Gate et Orcid, Scopus, ResearchID, LinkedIn...). Ces outils peuvent à la fois générer des classifications de thématiques des chercheurs et des publications et de mettre en relation les chercheurs. Mendeley, par exemple, permet de générer des bibliographies automatiques, de gérer une bibliographie en collaboration avec d'autres chercheurs, de trier de la bibliographie et d'être en relation avec sa communauté

Des mécanismes d'alertes de publications dans les domaines de travail en cours sont généralement demandés, sauf dans les communautés, comme celles de la physique, où le nombre des publiants est connu, limité et relativement stable (le chiffre de +5%/an d'accroissement des publications a été cité).

Enfin, une démarche en cours grandissante met l'accent sur les métriques "alternatives" (Altmetrics), qui s'appuie sur les pratiques des chercheurs et les analyses (préférences et opinions, notamment). Il y a place également pour les analyses heuristiques et leurs développements en cours (Méta et Hyper-heuristiques nées dans les applications industrielles de la recherche).

- Développer une analyse comparative des outils de valorisation des publications, en évaluer les risques et les avantages, et en partager les usages ainsi que la "mise sur étagère" d'outils communs et/ou d'outils partagés par communautés de recherche,
- Relier cette évaluation à la prospective IST (principe d'action 1),

- Développer les pratiques et usages associés à de nouvelles métriques (recommandation du Cléo citée plus haut). Promouvoir le développement en cours des "Altmetrics".

---

## LA REFONTE DE L'OFFRE : UN INIST RENOVE, UN PROJET NATIONAL

En relation étroite avec la DIST, en prise avec tous les besoins du CNRS et de la recherche, l'INIST a pris toute la mesure de son positionnement dans le Schéma d'Orientation stratégique de l'IST du CNRS, et, à travers cette démarche, de son positionnement dans toute la sphère de l'ESR.

Ces perspectives s'inscrivent dans un ancrage territorial fort, celui d'un pôle d'excellence IST en Lorraine. L'INIST, en association avec des partenaires universitaires dont en particulier l'Université de Lorraine, dispose d'atouts pour développer le pôle d'excellence en IST qui s'impose au niveau national. Cette conviction s'appuie sur les compétences qu'a développées l'INIST mais aussi la présence de recherche de qualité sur le traitement automatique des langues, tant au sein de l'ATILF que du LORIA (plus spécialement dans son département « Traitement automatique des langues et des connaissances »). La présence d'un laboratoire de recherche de qualité en sciences de l'information et de la communication, le CREM, complète ces ressources.

La rénovation de l'offre, préparée par un Séminaire stratégique où l'INIST a confronté ses analyses avec les Instituts et avec la DIST, est approchée de façon pragmatique, par type de prestations.

---

### 4.4 LE REAMENAGEMENT DES PORTAILS

L'INIST est partie prenante active de la réflexion engagée par la DIST sur les conditions et modalités d'un réaménagement de l'offre documentaire des portails existants. Cette réflexion, qui doit se faire avec les instituts, doit déboucher à brève échéance sur de nouvelles propositions, construites dans le cadre d'un Plan d'Actions défini pour faire bénéficier l'ensemble le plus large possible de chercheurs, d'une offre correspondant à leurs besoins en tenant compte des besoins spécifiques de chaque communauté. La présence de représentants de tout l'ESR dans le Plan d'Actions devrait donner à cette réflexion le cadre de décisions optimisées pour le début de 2014.

- Mettre en place le groupe de pilotage d'un réaménagement des portails, comportant la participation effective des Instituts et des partenaires du CNRS dans la recherche (Universités, EPST...),
- Définir, préalablement avec la DSFIM et la D2AI, les bases budgétaires à retenir pour encadrer les évolutions possibles.

---

### 4.5 LES SERVICES D'APPUI A LA RECHERCHE

L'INIST offre de nombreux services de qualité reconnue, certains sont à aménager et d'autres à promouvoir. Cet ensemble de services est un atout décisif dans la rénovation de l'offre de l'INIST, et constitue une demande forte des Instituts (ingénierie documentaire, curation de données, ingénierie de la langue, veille, études bibliométriques, traductions, formations, publication numérique, fourniture de documents, gestion et valorisation des données de la recherche).

#### 1. LA FOURNITURE DE DOCUMENT

La Fourniture de document (permet de fournir des articles rares et difficilement identifiables et localisables. Le service de fourniture de document de l'INIST permet aujourd'hui d'atteindre plus de 96% de satisfaction des commandes (notamment grâce à un réseau de 280 bibliothèques partenaires).

- Sous des conditions nouvelles à définir, cette activité portant sur des demandes de documents rares ou nécessitant des compléments d'identification pourrait apporter une véritable valeur ajoutée au dispositif académique géré par l'ABES. L'intégration de ressources en accès libre devrait également être un levier de

signalement renforcé. BSN 8 a retenu parmi ses priorités une meilleure articulation entre les ressources des Cadist et la plateforme RefDoc. Des études sont en cours.

## 2. LES BASES DE DONNEES BIBLIOGRAPHIQUES

Dans le cadre national de la Bibliothèque Scientifique Numérique, le périmètre des bases de données bibliographiques de l'INIST intègre le troisième segment (BSN3) dédié au signalement qui mentionne : «D'autres dispositifs de signalement se sont progressivement mis en place, répondant à des besoins spécifiques ou s'adressant à des communautés ciblées, notamment (...) les bases de données par grands domaines disciplinaires (Francis, Pascal de l'INIST) pour la valorisation de la production scientifique... ». Ces perspectives sont à construire au plus près des besoins des communautés, des unités et de la politique IST du CNRS.

Les objectifs sont :

- Etablir un lien entre les bases de données et le texte intégral et les données de la recherche,
- Produire un schéma de repositionnement de l'offre de l'INIST, tenant compte des acquis actuels comme des développements en cours de la demande,
- Construire le schéma de formation indispensable à l'accompagnement de ces évolutions.

## 3. L'INGENIERIE DE NEGOCIATION

L'ingénierie de négociation d'achat de ressources numériques est une activité dans laquelle l'INIST bénéficie d'une expérience acquise depuis presque 10 ans dans les négociations thématiques (pour les portails) ainsi que de celle acquise dans les négociations nationales (au côté de Couperin, des EPST et de l'ABES) dans lesquelles l'INIST représente le CNRS. Ces négociations concernent annuellement plus de 50 éditeurs (70 bouquets de revues ou bases de données) afin d'obtenir des baisses tarifaires, la pérennité des accès et la mise à disposition de statistiques d'usage.

Sur cette base, l'INIST peut apporter :

- Une réflexion et un conseil aux Instituts autour du besoin du chercheur. (l'InSHS souhaite un service au plus près des besoins, l'INSIS n'exprime pas de besoins (à part IEEE), l'IN2P3 n'a pas de besoins identifiés)
- L'Etude du coût et des différents schémas possibles,
- Une analyse, prévue plus haut, de l'évolution des portails ou du système portail, en relation avec la DIST/DSFIM/DDAI et COUPERIN, avec l'élimination des doublons internes au CNRS, la mise en évidence d'opportunité de commandes groupées, un travail de communication pour informer les laboratoires des possibilités et de l'existant et de son évolution, et enfin une réduction de la double irrigation avec les universités.

## 4. PRESTATIONS A LA DEMANDE

L'INIST peut également compléter l'offre sous forme de prestations réalisées à la demande des chercheurs par un accompagnement des professionnels de l'information présents dans les laboratoires, en termes de conseil méthodologique et technologique (recommandation d'outils, aide à leurs usages, alertes).

## 5. LA VEILLE STRATEGIQUE SUR L'IST

L'INIST offre un dispositif de veille avec notamment deux sites de qualité : <http://www.donneesdelarecherche.fr> et <http://openaccess.inist.fr>.

Les évolutions attendues visent à situer ce dispositif au cœur de la démarche IST du CNRS et de ses partenaires, pour en faire l'outil de toutes les évolutions en cours, vers des "connaissances mieux partagées". Cette évolution suppose de :

- Renforcer la visibilité : contact est pris avec la DSI pour accroître la visibilité via les moteurs de recherche,
- Intégrer des études scientifiques sur les métriques utilisées, des alertes sur les outils à disposition des chercheurs, des expertises sur la connaissance avancée des contenus d'une revue ou conférence,

- Rédiger des Fiches récapitulatives FAQ,
- Créer un réseau de correspondants pour alimenter les 2 sites Données de la recherche et Libre accès, communiquer pour les faire connaître, traduire des textes de référence, restructurer les sites pour mettre en relief les domaines scientifiques.

## 6. LA TERMINOLOGIE SCIENTIFIQUE

Issue initialement des activités liées à la mise au point des bases de données et leur indexation, la terminologie scientifique constitue un réservoir à mettre à disposition de l'ensemble de la communauté de l'ESR. Depuis 2005, le portail TermSciences (plateforme malheureusement demeurée figée) valorise et rend accessible les vocabulaires de l'INIST, de différents partenaires de la recherche et des institutions publiques (INRA, IRSTEA, Inserm) en permettant leur alignement terminologique

L'INIST est par ailleurs impliqué dans des projets de R&D et investissement d'avenir intégrant une part importante de fouille de textes pour proposer un accès élaboré aux connaissances scientifiques décrites dans les publications que l'INIST met à disposition de la communauté. Les projets actuels sont :

- ISTEEX : enrichissement des textes par reconnaissance de termes, entités nommées, citations, nécessaires au développement de services à valeur ajoutée,
- Quareo : annotation de textes et développement d'outils d'analyse sémantique de textes scientifiques,
- ANR TermITH (Terminologie et Indexation de Textes en sciences Humaines) : accès à l'information des documents numériques par le biais d'une indexation fondée sur les termes qu'ils contiennent, ce qui suppose reconnaissance, désambiguïsation et analyse des termes,
- UBIK : détection et caractérisation de vocabulaires transdisciplinaires, approche complémentaire des travaux sur les terminologies.

L'INIST doit développer une expertise forte en collaborant avec les meilleures équipes de recherche du domaine dans le cadre de projets portant sur le développement et la gestion de ressources terminologiques, le traitement automatique des langues, l'extraction et la gestion de connaissances.

Les applications de ces techniques d'ingénierie de la langue sont multiples dans les services que l'INIST peut apporter à la communauté. Un volet des Plans d'Action de l'INIST se développera avec l'objectif de décrire les prestations offertes selon les buts suivants :

- Faciliter un accès précis à l'information recherchée via des fonctionnalités de recherche sémantique,
- Améliorer la visibilité et le référencement des productions et données de recherche, via une annotation riche des contenus,
- Faciliter l'analyse du contenu des productions scientifiques et le traitement de gros volumes, par exemple pour des besoins de veille ou de pilotage scientifique.

## 7. LA RECHERCHE EN IST

Ces recherches s'organisent a

tour des trois axes : les mathématiques appliquées, les techniques symboliques et numériques de l'intelligence artificielle et le Traitement Automatique du Langage (TAL).

Les travaux ont porté sur l'analyse thématique de corpus scientifiques, la détection de thématiques et technologies émergentes, l'aide à la gestion d'appels à projets, notamment par le développement de nouveaux indicateurs bibliométriques



et infométriques mettant en œuvre et combinant des approches numériques de classification statiques puis incrémentales avec des approches de traitement automatique de textes.

- Définir un projet de partage des résultats et de valorisation de la recherche.

#### PRINCIPE D'ACTION 5 : PARTAGER LES RESSOURCES COMMUNES HAL, PERSEE

HAL est le dispositif national d'archives ouvertes de publications : sur cette infrastructure se concentrent toutes les attentes envers le développement d'une recherche publique partagée.

---

#### 5.1 L'EVOLUTION DES SERVICES OFFERTS EN OPEN ACCESS

Le CCSD vise à définir des services d'Open Process, qui peuvent être par ailleurs valorisés par une incitation forte au dépôt. Il faut développer fortement l'intérêt du chercheur pour le dépôt dans HAL. Pour ce faire, il convient de :

- Rendre l'outil attractif, et invisible.

Le dépôt est facilité (extraction facile quel que soit le choix de format, interaction avec les plateformes usuelles des chercheurs PubMedCentral, ArXiv, Openaire plus, récupération des métadonnées à partir des réservoirs existants (logiciels bibliographiques ou grandes bases institutionnelles, d'éditeurs,...), lier une publication à ses sources, ses données, ses matériaux...

- De nouveaux services d'Open Process

Des services sont prévus dans ce sens par le CCSD: Développer les procédures d'interopérabilité, Améliorer les procédures de dépôt, Améliorer les services aux chercheurs (statistiques, CV bibliographie). Il conviendra sans nul doute de donner vigueur à cette démarche interne positive en y adossant une réflexion d'ensemble nationale sur les services d'Open Access, que le CNRS pourrait mener avec tous ses partenaires dans le cadre du Plan d'Action IST. Le CCSD est aussi un acteur privilégié dans des projets de modèles d'Open Access innovants comme celui des épi-journaux

- Développer une incitation forte de la tutelle

Une réflexion nationale est nécessaire pour définir le bon niveau d'exigence : c'est une démarche qui reste à construire entièrement, au niveau de tout l'ESR comme au niveau interne du CNRS.

---

#### LE POSITIONNEMENT DU CCSD

Le Centre de communication scientifique technique directe (CCSD) a été créé en 2000 pour permettre la communication directe entre chercheurs. Par le dépôt de sa communication scientifique sous forme d'article dans l'archive ouverte verte HAL, le chercheur met l'IST à disposition de la communauté globale. Le CCSD a évolué d'UPS en UMS et le dispositif d'archive institutionnelle HAL (hyper article en ligne) a été redéfini dans la convention inter-établissements signée le 2 avril 2013 par un ensemble de partenaires représentatifs de tout l'ESR. Le CCSD a comme fonction de mettre en œuvre la stratégie du CNRS en termes d'archives ouvertes et son expertise à disposition des établissements de l'ESR. Sa mission principale est centrée sur le développement, l'exploitation et l'évolution des archives ouvertes HAL, TEL ou encore MédiHAL. Le CCSD entretient de forts partenariats avec l'ABES, le CINES, Huma-Num, Le Cleo, l'Inria.

- Les projets du CCSD sont conçus et menés dans l'optique de favoriser le développement de l'Open Access. La France a fait le choix de soutenir fortement la voie verte et HAL répond à cette politique en respectant notamment les recommandations du BOAI dans la mise en œuvre de ses archives ouvertes.
- HAL accueille une large typologie de documents en dépôt, correspondant aux usages des scientifiques, dépôts de prépublications ou de documents édités.

- Pour garantir un niveau scientifique minimum, et la qualité des métadonnées requises, le CCSD assure une modération et une vérification pour chaque dépôt, notamment des droits de mise en ligne des articles édités.
- HAL est régulièrement utilisé comme corpus de recherche. Là aussi la mise en place d'outils de fouille de données et de reporting favorisera ces recherches.

---

## LE POSITIONNEMENT DE PERSEE

Initié en 2003 par le ministère de l'enseignement supérieur et de la recherche et confirmé en 2007, le programme national Persée a pour missions la numérisation, le traitement documentaire, la diffusion et l'archivage pérenne de collections rétrospectives de revues et de publications scientifiques francophones (publications en série, actes de colloques) en sciences humaines et sociales, dans une logique de diffusion gratuite et non exclusive. Le changement de portage institutionnel et la mise en place d'une UMS réunissant des acteurs nationaux et locaux offre l'opportunité de franchir une nouvelle étape pour le programme Persée et de définir une nouvelle stratégie de développement à court et moyen terme, conciliant des missions de portée nationale et un enracinement fort dans le pôle lyonnais d'enseignement supérieur.

Persée est fortement impliquée au sein de l'IR documentaire BSN (BSN5) et sera amenée à participer à l'IR ColLEX (Collections d'excellence...).

Persée s'attache à développer les Interactions avec les opérateurs français d'IST. L'UMS partage des valeurs communes avec le CCSD, le Cléo et le programme NUMDAM de la cellule MathDoc, notamment l'implication dans le mouvement de l'Open Access et par conséquent, la diffusion ouverte de documents scientifiques ; il existe une complémentarité forte entre l'UMS Persée et la TGIR Huma-Num. Cela pourrait prendre la forme :

- d'échange d'expériences dans la modélisation des données et des métadonnées, les méthodes d'indexation et l'identification des documents,
- de mise en œuvre de passerelles entre les différentes plateformes (NUMDAM, CCSD, Cléo) par le biais de rebonds entre les documents (liste d'autorités auteurs et référencement croisé).

---

## 5.2 LES SERVICES OFFERTS PAR PERSEE

L'UMS Persée a pour objectif de créer des corpus scientifiques numériques pérennes, de les enrichir, de les diffuser et d'y associer des outils de signalement, de recherche et d'exploitation. Cette offre répond à un usage de recherche et elle couvre tous les domaines scientifiques. A ce titre, l'Unité s'investit notamment dans :

- Le développement d'une offre structurée, pérenne, en Open Access de documents numériques à caractère scientifique,
- L'Innovation en matière d'outils et de méthodes et accompagner l'émergence de nouveaux usages liés au numérique.

L'Unité remplit des missions nationales de consultation, de formation et d'appui à des projets de constitution de corpus scientifiques numériques au service de la recherche.

- Devenir un pôle national de référence et de compétences.

Elle est moteur d'innovation technologique par :

- La conception et le développement d'un système générique destiné à héberger la production, le traitement, la diffusion et l'archivage de différents types de ressources scientifiques, en se fondant sur l'expérience acquise dans le cadre de la gestion du programme Persée ( [www.persee.fr](http://www.persee.fr) ),
- La facilitation des utilisations transversales par la mise en œuvre sur un socle technologique commun : navigation ou recherche multi-corpus, enrichissement mutuel des données présentes, approche transdisciplinaire des corpus,

- Le développement d'outils d'indexation et de recherche,
- Collaboration avec des équipes de recherche autour :
  - des sciences de l'information et ingénierie des connaissances,
  - du web sémantique,
  - de la modélisation des systèmes sociaux à partir de grandes masses de données.

Mais elle également moteur d'innovation juridique car elle opère dans un cadre juridique complexe à anticiper et à faire évoluer. En effet, l'UMS ne limite pas son action aux corpus libres de droit et prend en charge des documents qui peuvent être soumis au droit d'auteur. Les aspects juridiques occupent en conséquence une place centrale. Il convient non seulement de connaître les ayants droits sur les documents et les droits dont ils disposent mais également d'organiser les cessions de droit nécessaires et d'assurer les conditions d'une diffusion en Open Access. Cela se traduit par :

- Des études de faisabilité préalables,
- Une réflexion sur l'évolution du droit d'auteur.

## PRINCIPE D'ACTION 6 : VALORISER ET PARTAGER LES DONNEES DE LA RECHERCHE

### UN PERIMETRE ETENDU

Avant même d'évoquer les besoins et la demande des Instituts, la question posée à l'IST par la prolifération des données appelle des réponses à plusieurs niveaux, avec autant de démarches à construire. BSN10 va permettre d'établir des chartes communes pour tous. Cette création précède de peu les préconisations de l'Europe pour 2014 qui après celles consacrées aux publications en 2013 se projettent sur les données et leur accès.

Ces instances à différents niveaux vont permettre d'établir des règles communes qui commencent à être suivies et connues des disciplines les plus avancées dans le web des données.

Les sujets génériques du partage des données de la recherche, peuvent être décrits dans l'ordre ci-dessous :

- Quelles données partager ?
- Définir le cycle de vie des données,
- Garantir les liens entre publications et données,
- Sensibiliser les chercheurs à la gestion de leurs productions scientifiques (y compris éthique),
- Mettre en place des plans de gestion de données,
- Participer à la définition et l'application des normes et standards de description, citation, qualité, interopérabilité,...
- Informer sur la propriété intellectuelle applicable aux données,
- Assurer l'archivage pérenne des données selon les besoins des communautés,
- Installer et maintenir les infrastructures techniques adéquates,
- Prendre en compte la gestion des données dans le processus d'évaluation des chercheurs,
- Engager une gestion des compétences et des plans de formations des personnels de soutien à la recherche.

### UNE DEMANDE GENERALE FORTE DES INSTITUTS

L'INEE est partenaire de grandes bases de données internationales (GBIF, WOORS) et développe une politique systématique de valorisation. SCAR-MarBIN, par exemple, est un portail de bases de données interopérables, auquel le CNRS contribue en fournissant des données et la Belgique offre l'Infrastructure. Zookeys et Phytokeys sont des journaux qui éditent des articles automatiquement à partir de bases de données.

Pour l'INEE, il faut définir une incitation du CNRS à l'ouverture des données avec une obligation à citer, tout en admettant qu'il existe bien sûr des données sensibles à ne pas diffuser. Les données "sont les alliées" des publications, qu'elles complètent systématiquement pour le test et la contre-épreuve du travail scientifique (INEE, INSIS, INSB, INP).

IN2P3 ne publie pas les données brutes (illisibles pour des tiers) mais les Data management plans sont connus et utilisés. Le TGIR SOLEIL expose également un besoin, au-delà de la gestion des données, sur les règles de stockage à long terme.

INSB avec la bioinformatique, est partenaire d'infrastructures nationales (avec l'Inria), pour les ressources biologiques et biomédicales (biobanques). L'accès aux données est familier (séquençage et structure des protéines). L'INS2I souligne la question de la curation des données : c'est une question importante dont le traitement est délicat.

L'INSU produit des bases de données et est partenaire de bases de données internationales (Astronomie, sismologie, Géochimie...). Il souhaite à ce titre la valorisation des publications liées à ces bases de données. D'autres bases nécessitent une sécurisation dans le temps des données et une valorisation de celles-ci.

L'INSHS souligne l'importance d'une feuille de route CNRS sur le sujet.

Pour l'InSHS, il est impératif de sauvegarder les bases de données produites dans les laboratoires: l'Institut voudrait développer un projet de coopération avec l'INIST pour répondre aux besoins des laboratoires en matière de sauvegarde et d'interopérabilité des bases de données et ainsi faciliter à terme la diffusion des données de la recherche. Il paraît nécessaire de créer des triplets documentaliste, informaticien et chercheur destinés à aider les laboratoires à pérenniser les bases de données SHS : c'est une demande récurrente des unités ; cela doit être un de nos objectifs prioritaires..

Enfin, pour l'INSHS il faut renforcer les infrastructures : Huma-Num, Progedo-réseau Quételet, programmes d'excellence (Bibliissima <http://www.bibliissima-condorcet.fr/>) en stabilisant le rôle du réseau des MSH. Il faut également accompagner l'évolution des personnels d'appui en IST dans la gestion des données de la recherche.

L'INSMI utilise peu de bases de données généralistes comme WoS ou Scopus. En revanche, l'Institut utilise quotidiennement les deux bases dédiées MathSciNet (alias Math. Reviews, éditée par la société mathématique américaine AMS) et zbMATH (alias Zentralblatt, produit par le FIZ et commercialisé par Springer). Il existe des liens anciens entre la France et le zbMATH, dont la première version web a été développée par la cellule Mathdoc. De façon générale, il y a la volonté de soutenir cette base européenne pour limiter l'hégémonie (et les tarifs) de la base américaine. On pourrait imaginer utiliser cette base dans le portail math pour couvrir toutes les recherches documentaires sur la période 1868-2012 sur laquelle elle est quasiment exhaustive. Au-delà de la réflexion sur une conservation des supports papier, citées dans les objectifs, la communauté est consciente des enjeux et limites en matière d'archivage pérenne des ressources numériques de plus en plus nombreuses qu'elles soient issues de programmes de numérisation ou de documents natifs électroniques.

Un exemple probant de cette problématique est que les collections numériques produites par la cellule Mathdoc (NUMDAM, CEDRAM) ou acquises par le RNBM (contrats nationaux Springer, EMS-PH, SMF, etc.) ne sont pas archivées actuellement dans un système pérenne. Le modèle d'affaire du CINES est inadapté : un tel archivage devrait être pris en charge par le MESR pour les collections labellisées BSN comme cela a été recommandé par les groupes BSN 5 et 6. La mise en place d'un réseau CLOCKSS avec les partenaires de EuDML initiative sera étudiée en 2014.

- Il y a donc aujourd'hui un important besoin d'archivage pérenne, de recensement de l'existant et par la suite de pouvoir faire des choix quant à ce qu'il faut garder ou pas. Les universités et les instituts de recherche sont intéressés par la gestion et l'exposition des bases de données produites dans leurs établissements, mais ne relevant pas du champ d'application d'entrepôts thématiques ou institutionnels.
- Ce spectre étant très large, des coopérations sont à valoriser, par exemple avec Huma-Num (très grande infrastructure dédiée à la recherche en SHS) travaillant sur l'objet des humanités numériques ou l'UMS 3468 « Bases de données Biodiversité, Ecologie, Environnements Sociétés (BBEES) ».
- L'INIST est en capacité d'intervenir dans les formations initiales ou continues vers les communautés (école doctorale et chercheurs) et vers les professionnels de l'IST.
- Il y a nécessité de mener un recensement national partagé des besoins.

Le partage des données de la recherche s'inscrit aujourd'hui dans la dynamique générale de l'Open Data, visant à libéraliser l'accès aux données gouvernementales, porté en France par ETALAB, avec lequel un contact préliminaire est pris par la DIST.

La prise de conscience de l'intérêt pour le développement des connaissances du partage des données de la recherche, notamment lorsque la production de celles-ci est financé sur des fonds publics, est déjà ancienne.

La déclaration de Berlin ( 2003), les principes OCDE d'accès aux données de recherche (2006), les principes des conseils de recherche britanniques (2011), les recommandations de la Commission européennes (2012) suite au rapport « Riding the Wave » (2010), les directives de l'OSTP vers les agences fédérales de recherche américaines (2013), sont autant de recommandations, principes voire injonctions à partager, rendre accessibles les productions scientifiques de recherche, tout comme la directive européenne 2007/2/CE du 14 mars 2007 (INSPIRE), qui vise à établir une infrastructure d'information géographique dans la Communauté européenne pour favoriser la protection de l'environnement. Son application aux données de la recherche fait l'objet de soins attentifs de la part de l'INEE et de l'INSU.

---

## DES REGLES ET CONDITIONS COMMUNES DE VALORISATION

### DES CONDITIONS COMMUNES INCONTOURNABLES

Deux ensembles de conditions se dégagent des demandes : les premières sont institutionnelles, les autres individuelles.

#### 1 CONDITIONS LIEES AU SOUTIEN INSTITUTIONNEL

- Partager un texte fédérateur à l'échelle du CNRS, traduisant un engagement sans ambiguïté en faveur des données de recherche,
- Définir et faire connaître la politique de l'organisme concernant les données de la recherche.

Ce principe d'action correspond à la nécessité institutionnelle de disposer d'un texte de politique générale concernant les données définissant les modalités de mise en ligne et de partage, précisant les cadres juridiques et les modèles économiques afférents.

- Cette politique, nourrie d'une réflexion inter-organisme, doit pouvoir être déclinée dans les différents instituts et prévoir à ce niveau des mécanismes d'incitation et de régulation. Elle devra traiter du sujet de la visibilité des bases de données produites, de l'avancement des chercheurs et personnels impliqués dans leur maintenance opérationnelle.

#### 2 CONDITIONS LIEES A L'ADHESION INDIVIDUELLE

- Les chercheurs doivent être sensibilisés (par des pairs) à l'idée de partage entre eux et vers la société de leurs productions scientifiques, données de la recherche comprises.
- Pour favoriser ce changement et obtenir l'adhésion des chercheurs et des communautés scientifiques, il est indispensable de développer des règles de reconnaissance pour les activités de partage, de curation et de diffusion des données de recherche.

---

## DES CONDITIONS D'ENVIRONNEMENT A REUNIR

#### 1 CONDITIONS LIEES A LA FORMATION

- Développer à l'échelle de l'organisme les compétences attachées à la gestion des données de la recherche.

La production généralisée de données numériques induit de facto des besoins importants de formation continue permettant la maîtrise des descriptions, stockage, traitement, analyse,...

Ces formations ne se limitent pas à la connaissance des outils technologiques pour accomplir ces tâches mais englobent également les aspects méthodologiques, juridiques et organisationnels nécessaires au montage des plans de gestion de données et fonctionnement des infrastructures.

Pour partie, ces formations devront être communes aux trois corps de métiers qui concourent à la gestion efficace des infrastructures : scientifique, informaticien et professionnel de l'information. La description des données met en jeu aussi des connaissances du champ scientifique.

## 2 CONDITIONS LIEES A LA GESTION DES DONNEES

- Mettre en place des plans de gestion de données dès la conception des projets de recherche.

Ce principe d'action vise à ce que, dès la genèse des projets, les chercheurs prennent en compte les aspects des modalités de stockage, de description, de mise à disposition et d'archivage des données qui seront amenées à être collectées et traitées. L'établissement d'un plan de gestion de données est de plus en plus fréquemment demandé dans les appels à projet. Il s'appuie sur la prise en compte du cycle de vie des données et décrit les règles applicables en termes de normes de métadonnées, format des bases de données, méthode et sécurité d'accès, durée d'archivage, ainsi que les coûts associés.

- Une mention particulière doit être apportée aux données venant en appui des publications et qui doivent à ce titre rester disponibles et accessibles le plus largement possible. Les personnels de soutien, informaticiens et professionnels de l'IST sont sollicités pour ces tâches.

## 3 CONDITIONS LIEES AUX INFRASTRUCTURES

Pour les disciplines qui relèvent des Big Data (HEP par exemple), les infrastructures de calcul assurent les fonctions de base de gestion des données (stockage, sauvegarde), mais rarement très au-delà de la vie des projets. Une infrastructure de données se doit d'inclure dans ses services les dimensions archivage pérenne, règles de partage et de réutilisation, gestion des accès, ...

- Identifier les structures existantes remplissant déjà certaines fonctions.
- Définir et opérer des infrastructures de données aptes à répondre aux besoins de l'organisme.
- Disposer de fonctionnalités de stockage, de mise en ligne, d'archivage à long terme, permettant d'assurer l'intégrité et le partage des données au cours et après la vie d'un projet, tel est l'enjeu.
- Si les projets Big Data s'appuient sur des infrastructures de traitement assurant également le stockage des données, il n'en est pas de même des nombreux projets «small data» qui présentent des solutions de gestion de données moins sûres et moins pérennes.
- Il s'agit donc de mettre en place une offre de services, en liaison ou non avec l'offre de calcul, s'appuyant sur les moyens des laboratoires, les centres de type CC-IN2P3, CINES, mais aussi INIST et cloud CNRS.

## 4 CONDITIONS LIEES A LA RECHERCHE EN IST ET A L'ANALYSE DES CAPACITES DE CALCUL

- Progresser dans l'ouverture, le partage et l'usage des données de la recherche nécessite de disposer d'une recherche appliquée active sur laquelle une **ingénierie de la connaissance** pourra s'élaborer. Le volet « développement d'outils avancés » du projet ISTE en est l'exemple le plus frappant.
- Dans le même sens, il importe de demander au CNRS de développer une approche prévisionnelle des capacités de calcul disponibles en interne (COCIN, DSI), RENATER, et également en relation avec le GENCI notamment (600 projets soumis actuellement).

## PRINCIPE D'ACTION 7 : IDENTIFIER ET ANALYSER LES PUBLICATIONS

Disposer d'une politique cohérente d'affiliation à laquelle travaille le SAPPs (Service d'appui à la politique et à la prospective scientifiques) est un facteur majeur d'optimisation pour toutes les étapes de construction de l'action d'IST. Ce truisme n'empêche en rien les pratiques de présenter des différences souvent peu probantes au regard des enjeux. D'une façon générale, la réflexion interne du CNRS a montré l'ambition d'avancer résolument vers des positions assurant la claire traçabilité nationale et internationale de sa production scientifique, qui comprend en moyenne annuellement quelques 29 000 articles<sup>37</sup>.

---

### DES PRATIQUES DISJOINTES

Le recensement des publications est fait par différents moyens, par des acteurs variés, pour de multiples usages :

---

<sup>37</sup> <http://www.cnrs.fr/fr/pdf/RA2012/index.html#/1>

- Communication scientifique entre chercheurs (dépôt dans ArXiv par les chercheurs de l'IN2P3 ou de l'INSMI par exemple)
- Description de l'activité de recherche (saisie dans l'application RIBAC<sup>38</sup> par les chercheurs en SHS)
- Signalement, mise à disposition du texte en libre accès, valorisation et gestion bibliographique (HAL dans tous les domaines)
- Etudes bibliométriques et aide au pilotage (repérage dans le WoS par les bibliomètres du SAPPS et de l'INIST)
- Evaluation individuelle du chercheur CNRS (dépôt de sa liste dans l'application e-valuation)
- Evaluation de l'unité ou demande de financement (par différents moyens : dépôt de dossier, formulaire de saisie).

Pourtant, le CNRS ne dispose pas d'un recensement exhaustif de ses publications et les instituts ne disposent pas toujours d'une vision complète (l'INC par exemple, le fait observer) et verraient d'un bon œil la mise à disposition d'outils de gestion bibliographique aux chercheurs (c'est le cas de l'INEE par exemple).

Des indicateurs bibliométriques (répartition par institut, relations interdisciplinaires, part nationale ou régionale, collaborations internationales,...) sont produits par le SAPPSS et sont largement utilisés.

Il apparaît également la demande d'une exploitation plus ciblée, de la base issue du WoS, par repérage des unités CNRS dans les affiliations, exprimée par certains instituts (INSB par exemple). D'autres méthodes de mesure seraient également souhaitées car la France n'apparaît toujours à la place qu'elle devrait avoir (c'est le cas de l'IN2P3 qui publie peu mais des papiers de qualité). Dès à présent, des travaux d'analyses des publications ont été réalisés par plusieurs Instituts à partir du WoS (INEE, INP notamment) pour mieux connaître l'activité des unités ou détecter les collaborations entre unités par exemple avec les outils dont ils disposent. L'INSB, par exemple, exprime la volonté de mieux valoriser son activité de recherche, en exprimant l'idée que l'écriture des affiliations est une question centrale.

Des outils d'analyse ou des indicateurs de popularité ou citation, notamment, sont utilisés par les chercheurs (Google scholar, Publish or perish, Research Gate, H-Index du WoS,...). Une forte demande est exprimée dans ce sens. Il a déjà été préconisé à plusieurs reprises que, notamment dans les contrats entre le CNRS et ses partenaires, que la mention de CNRS est obligatoire: en pratique ce n'est pas toujours le cas.

---

## UNE CONVERGENCE A CONSTRUIRE

De nombreux dispositifs de référencement des publications par les chercheurs existent et pourtant de nouveaux outils sont en gestation alors que le problème organisationnel global reste en suspens. On constate une multiplication des outils d'analyse et des usages (Incites, Research gate, Google Scholar Citations ...) et une incitation à l'utilisation d'identifiants auteur.

Des alternatives aux indicateurs historiques tels que le Facteur d'impact apparaissent, et sont proposées à l'échelle internationale (cf. le mouvement Altmetrics<sup>39</sup> déjà cité).

Des chartes de signatures des publications et des mesures incitatives sont mises en place au sein des universités pour faciliter les repérages, améliorer la visibilité et les classements mondiaux.

Un savoir-faire existe au sein du CNRS (SAPPSS) en matière de référencement et analyse des publications (SAPPSS).

---

## UN CHEMIN VERS LA TRAÇABILITE DE L'IST

---

<sup>38</sup> Recueil d'Informations pour un oBservatoire des Activités de reCherche en SHS

<sup>39</sup> <http://en.wikipedia.org/wiki/Altmetrics>



- Clarifier, optimiser les dispositifs de recensement des publications utilisés au CNRS
- Evaluer les outils d'analyse des publications
- Analyser les usages des nouveaux outils proposés aux chercheurs
- Mener une réflexion nationale sur les "diligences minimales communes" en vue d'assurer une présentation homogène des identifiants de publication.

## PRINCIPE D'ACTION 8 : UNE STRATEGIE GLOBALE POUR LE LIBRE ACCES A LA SCIENCE PUBLIQUE

Aller vers le Libre-accès aux résultats et aux bénéfices de la recherche est un long chemin dans lequel des étapes majeures ont été accomplies et s'accomplissent au quotidien : aucun enjeu en ce début de siècle n'aura autant mobilisé la communauté internationale et les responsables de la recherche.

L'enjeu global est considérable aussi bien du point de vue du partage des bénéfices de la recherche que de celui des conditions de production de la science, entravée qu'elle est aujourd'hui par les multiples péages qu'imposent les entreprises qui contrôlent les divers stades de la publication d'un grand nombre de disciplines.

Le consensus est maintenant bien établi et les difficultés apparaissent : grâce aux efforts des pionniers parmi lesquels le CNRS figure clairement, le débat est passé des questions de principes à celles des modalités. Les modalités de passage à l'Open Access constituent donc une question stratégique centrale, dont les applications potentielles sont présentes dans toute l'action d'IST.

---

### LE PASSE

Par la Déclaration de Berlin, le CNRS s'engageait il y a dix ans dans la voie du libre-accès aux résultats de la recherche publique, position réaffirmée par les « *Position statement* » dans le cadre de ScienceEurope en 2013<sup>40</sup>. Dans le même sens, la Recommandation aux états membres de la Commission Européenne en 2012<sup>41</sup>, s'impose, tout comme la position ministérielle précitée de 2013<sup>42</sup> et les recommandations du G8<sup>43</sup>.

Depuis dix ans, les déclarations, initiatives, créations d'archives ouvertes et de journaux en *Open Access* se sont multipliés. L'INIST a mis en place un outil de veille fort complet, « *Libre accès à l'information scientifique et technique* »<sup>44</sup>, qui permet de suivre cette dynamique. Le CCSD, dès sa création, a toujours mené des actions de communication, formation et promotion de l'open access auprès des communautés professionnelles et scientifiques

---

### LE PRESENT

#### UNE CONVERGENCE GLOBALE

---

<sup>40</sup> [http://www.scienceeurope.org/uploads/Public%20documents%20and%20speeches/SE\\_OA\\_Pos\\_Statement.pdf](http://www.scienceeurope.org/uploads/Public%20documents%20and%20speeches/SE_OA_Pos_Statement.pdf)

<sup>41</sup> [http://ec.europa.eu/research/science-society/document\\_library/pdf\\_06/recommendation-access-and-preservation-scientific-information\\_fr.pdf](http://ec.europa.eu/research/science-society/document_library/pdf_06/recommendation-access-and-preservation-scientific-information_fr.pdf)

<sup>42</sup> <http://www.enseignementsup-recherche.gouv.fr/cid66992/discours-de-genevieve-fioraso-lors-des-5e-journees-open-access.html>

<sup>43</sup> <http://openaccess.inist.fr/?Le-G8-donne-sa-position-sur-le>

<sup>44</sup> <http://openaccess.inist.fr/>



Le libre accès, tout comme la recherche elle-même, est une question globale, traitée par ailleurs à de nombreux niveaux internationaux (Science Europe, Global Research Council, Knowledge Exchange COAR<sup>45</sup> etc), au niveau européen, au niveau national dans le cadre de la BSN. Au-delà des déclarations et des engagements, un nombre rapidement croissant d'institutions (agences de financement, universités, organismes voire gouvernements) définissent et mettent en œuvre de véritables politiques de libre accès. Cette progression paraît aujourd'hui irréversible.

## UNE INCERTITUDE SUR LE MODELE ECONOMIQUE

---

La convergence de principe n'exclut pas les divergences, avec une opposition portant, notamment sur le modèle d'auto-archivage (voie verte) et le modèle de publication dans des journaux « Open Access» (voie dorée).

Il y a en réalité trois modalités et non pas les deux présentées habituellement (verte et dorée) :

1. Archives ouvertes, vert
2. Doré avec frais de publication réglés par l'auteur ou son institution (dites APC, pour « *article processing charges* »)
3. Doré sans frais, nouvelles formes d'édition

## UNE CERTAINE CONFUSION, QUI REFLETE PLUTOT LA RICHESSE DU DEBAT

---

On confond trop souvent le dépôt d'une notice bibliographique dans une archive avec du libre accès.

Il y a illusion que le libre accès est déjà une réalité pour ceux qui sont satisfaits de l'offre des portails et oublient que quelqu'un a payé les abonnements pour eux. Des craintes s'expriment, concernant le risque de pillage en lien avec les enjeux de propriété industrielle et commerciale en raison du libre accès, alors qu'il ne s'agit que de publication, comme dans les revues accessibles sur abonnement.

Grâce à l'efficacité du lobbying des grands éditeurs privés, on confond presque toujours voie dorée, parfois même le libre accès lui-même, avec paiement par l'auteur. On s'inquiète des journaux prédateurs, qui n'offrent presque aucun service en prétendant faire évaluer la qualité des articles. Les instances d'évaluation se basent trop souvent sur la réputation de la revue (voir le facteur d'impact), ce qui entraîne un effet dissuasif pour le libre accès.

## DES RISQUES GENERIQUES

---

Plusieurs niveaux de risques sont à intégrer à l'analyse stratégique des choix nationaux en faveur du Libre-accès. Notamment ceux liés à :

- La préservation à long terme des archives,
- La prolifération des publications,
- La fragilisation des éditeurs « académiques »,
- La baisse générale de la qualité.

Ces quatre risques sont liés à l'IST numérisée en général, et non spécifiquement au libre accès, mais ce dernier ne peut que les aggraver.

## UN POSITIONNEMENT COMPLEXE MAIS UN "WORK IN PROGRESS"

---

Les voies d'évolution dans les trois domaines de l'Open Access présentent toutes un certain niveau d'incertitude, qui mérite d'être mis à plat.

ARCHIVES OUVERTES, VERT

---

<sup>45</sup> <http://www.scienceeurope.org/>, <http://www.globalresearchcouncil.org/>, <http://www.knowledge-exchange.info/>, <http://www.coar-repositories.org/>

Les Archives institutionnelles (dont la plateforme HAL et ses déclinaisons) comportent encore assez peu d'outils pour le chercheur dans son activité de recherche. Aujourd'hui seules les archives thématiques peuvent être l'outil du chercheur : elles s'adressent aux « communautés » (arXiv, RePec, PMC etc.), sans limitation institutionnelle ou territoriale.

En dehors de certaines disciplines (maths, physique des hautes énergies ou théorique et quelques autres...) les preprints ne sont pas considérés, mais seulement la production scientifique validée (articles principalement). Dans certaines communautés les chercheurs jugent avoir autre chose à faire que de déposer leurs publications dans une archive, surtout si des saisies ont déjà été faites par ailleurs (notamment archives disciplinaires). Donc, par défaut, la démarche "manque de bras".

Ainsi, fondamentalement, le mouvement des archives ouvertes semble irrésistible, même si les enthousiastes s'impatientent de la lenteur du changement : il y a encore trop peu de dépôts dans HAL, même si la pente est bonne, et que d'aucuns ressentent une insécurité juridique, quand bien même l'exemple des physiciens des hautes énergies ont-ils réussi sans jamais se poser de question avec la construction d'arXiv.

#### DORE AVEC FRAIS DE PUBLICATION

Avec l'Open Access doré "hybride", on fournit une rémunération supplémentaire aux éditeurs : les revues continuent d'être accessibles sur abonnement dont les coûts ne baissent pas.

Aux tarifs actuellement demandés par l'Édition, le pur doré est difficilement soutenable comme voie de généralisation du Libre-accès pour les organismes qui publient beaucoup. On ne sait évaluer les usages liés au "doré auteur-payeur", qui est représenté par le projet SCOAP<sup>3</sup> (IN2P3 et INP) et les journaux de l'EGU soutenus par l'INSU.

#### DORE SANS FRAIS, NOUVELLES FORMES D'ÉDITION

Des projets existent de nouvelles formes d'édition, déjà avancés ou mûrs, comme Open Edition (InSHS), ou naissants, comme le projet episciences.org avec les épijournaux en maths et informatique. Il en existe encore d'autres comme par exemple de nouvelles méthodes d'évaluation du type « open peer review ».

La publication conjointe avec celle des données de recherche se développe, voire la publication des données elles-mêmes, les « data journals ».

Les « mega-journals », s'ils relèvent aujourd'hui du « pur doré » n'en sont pas moins des initiatives originales dont l'arrivée peut bousculer le paysage de l'édition savante de manière importante. Il ne s'agit plus vraiment de journaux, mais de sortes d'archives ouvertes payantes et au contenu certifié.

---

## UNE STRATEGIE POUR L'AVENIR

L'avenir repose sur un ensemble de constats stratégiques qu'on peut résumer ainsi :

- Une stratégie du CNRS sur ces thèmes ne peut être que multi-partenaire et doit donc être bâtie avec les universités et les organismes de recherche.
- L'avenir ne peut être abordé qu'à partir d'une vision internationale, en liaison avec les réseaux existants ou à venir.
- Il n'y a pas de solution universelle. Il convient de rester ouvert à toutes initiatives et expérimentations, quand elles sont partagées et raisonnables : nouvelles formes de publication ou de communication, nouvelles formes d'édition (nouveaux « work flows », nouveaux modèles économiques, « par-delà vert et doré »).
- Une attention doit être portée aux éditeurs académiques et à l'édition scientifique publique, surtout en maths et en SHS.

## UNE REGULATION PARTAGEE: DYNAMISER LES PRATIQUES

### PRINCIPE D'ACTION 9 : PARTAGER L'ETHIQUE ET LES REGLES

#### UN BESOIN DE REGLES ETHIQUES POUR L'IST NUMERIQUE

#### UNE DEMANDE GENERALE DES INSTITUTS ET COMPOSANTES IST

La demande de repères est générale : avec qui partager les données ? Quelles sont les règles éthiques applicables à la conservation, à la divulgation ? Quelles doivent être les attitudes envers la publication à partir de données partagées ? Ces questions sont toutes vécues comme relevant d'une application de la Science ouverte, et relèvent aussi bien d'une politique d'établissement que d'une politique plus globale du partage de l'IST à l'heure numérique, comme l'observe l'INSU.

L'InSHS souligne l'importance et l'urgence de fournir un cadre éthique au travail sur les données relatives à la personne, aux opinions et aux préférences. Les SHS elles-mêmes peuvent contribuer à construire des références. La réflexion est étendue et pluridisciplinaire. Le TGIR Huma-Num a mis en place une démarche éthique interne sur les données ethnologiques et sociologiques. Comment traiter la question des traces laissées par les personnes (les chercheurs) dans les systèmes d'information ? Le Cléo souligne que les choix éthiques ne peuvent être abordés de façon unilatérale et sans débat. L'obligation de révéler les données devrait toutefois devenir la règle, afin de permettre la reproductibilité et la « confrontabilité » du travail de la science. L'INIST émet également une demande dans ces sens. RENATIS indique disposer d'un Code de déontologie rédigé par des professionnels des réseaux (2004-2005) à remettre à jour, diffuser et promouvoir auprès des professionnels de l'IST ; porter à la connaissance des directions d'Instituts, des directions des laboratoires, des chercheurs.

A l'INSB, on traite la question depuis longtemps, avec l'utilisation des cahiers de laboratoire assurant une traçabilité maîtrisée des données et des résultats, ainsi que de la publication. L'INSB dispose d'une Cellule Ethique qui veille au respect des réglementations en vigueur concernant les organismes génétiquement modifiés, les tissus humains, l'utilisation des animaux, y compris les primates non-humains et toute recherche sur l'homme. Il pourra être intéressant de partager cette expérience avec les autres Instituts.

Pour le TGIR SOLEIL, "on est au milieu du gué", dans un environnement très complexe, des initiatives se prennent au niveau européen ; les exigences sont variées et parfois contradictoires. On a des acteurs industriels pour lesquels (fourniture de données stockées sur des durées longues) il faudrait cadrer la dimension des responsabilités publiques. A partir de quand les données obtenues doivent devenir publiques et comment ? Les questions de SOLEIL sont « géologiques » : les questions techniques et juridiques doivent répondre à des règles pas seulement françaises mais européennes. Une Charte de « déontologie » paraît intéressante, au regard notamment du plagiat, des données erronées.

L'IN2P3 ne signale pas de problèmes d'accès à ses données, illisibles le plus souvent hors des contextes où celles-ci sont stockées.

Pour l'INSIS, en matière de réalité virtuelle, il faut pouvoir guider les expérimentations, comme c'est la règle pour le vivant.

L'INC souligne que les règles doivent être posées au bon niveau, qu'il faut chercher avant tout l'adhésion. Quel statut donner aux données « brutes » ? Aux données « interprétées » ? Quelle est la frontière ? La question se pose avec la caractérisation d'échantillons et la part que peut y laisser la trace laissée par l'outil d'analyse....

Quels critères éthiques doivent guider les choix de mise à disposition des données publiques ? L'INEE fait observer que la frontière est difficile à établir : les recherches sur contrat peuvent dans leur clauses, prohiber la mise à disposition des résultats et/ou des données : quels critères appliquer au regard de la libre disposition des résultats de la recherche publique ?

Pour l'accès aux données qui contiennent des informations de vie privée, les questions d'éthique sont posées, rappelle l'INS2I. Allistène, l'Alliance des sciences et technologies du numérique possède un comité d'éthique qui a été mis en place il y a un an (<http://www.inria.fr/actualite/actualites-inria/ethique-technologies-numeriques>). Il est présidé par Max Dauchet et Michèle Leduc du Comets y participe. L'utilisation des données de recherche fait partie des questions dont est saisi ce comité.

L'INSU, en particulier en astronomie, a une pratique très positive de l'accès ouvert aux données, tempéré par une période propriétaire – souvent un an – qui permet aux chercheurs ayant obtenu du temps d'observation sur appel d'offre d'étudier les données avant la mise à disposition de tous.

La question appelle un traitement systémique puisque travailler sur des données peut permettre d'en extraire des informations qui, à leur tour, peuvent caractériser les personnes. L'ordre de l'identité est ainsi bouleversé : il va des données à la personne. L'identité numérique peut se construire comme un résultat et non comme une donnée, telle que celle de l'état-civil.

Une charte de l'IST est recommandée par l'INSMI, en s'appuyant sur le travail fait au sein des Alliances, de l'INRA, du Comité d'éthique. La circulation des informations et le partage des connaissances deviennent la règle. L'indépendance du chercheur doit être affirmée, tout comme le refus d'un paiement pour l'accès aux résultats (réserve de principe envers les « Article Processing Charges »). Quelle est l'éthique applicable aux grands réservoirs de données publiques comme Hal et arXiv ? Comment fixer la paternité et la propriété des données qui y sont insérées ? Il faut par ailleurs des règles plus larges qu'Open Access. Faut-il définir un périmètre plus large : Open Science ?

ETALAB<sup>46</sup>, le site des données publiques (circulaire du Premier ministre du 17 septembre 2013), va nous amener à donner une réponse globale à ces questions sous la forme de la communication des données publiques de la recherche.

## UNE CHARTE D'ETHIQUE

---

Le besoin d'une Charte d'Ethique se décline à plusieurs niveaux :

- Mettre en place une réflexion sur les principes à décliner pour l'IST numérique,
- Décliner ces principes en fonction des objets d'IST concernés (données, publications) et des domaines d'application,
- Examiner les besoins de Chartes IST spécifiques par domaine ou projets de recherche en vue d'une contribution au dispositif public ETALAB,
- Mettre à contribution les pratiques existantes en France (Alliances, Organismes de recherche), comme au plan international (Europe, Grands organismes étrangers....).

---

<sup>46</sup> L'idée a fait son chemin : l'ANDS (Australian National Data Service) a entrepris de faire circuler et mettre en synergie les résultats de la recherche et les données de la société.

## LE DROIT NUMERIQUE APPLICABLE A L'IST

---

### LES BESOINS DES INSTITUTS

---

La demande de clarification est, là aussi, presque générale.

L'INEE souligne le besoin d'une élaboration partagée du cadre juridique, afin qu'il corresponde aux besoins des chercheurs et qu'il fournisse un outil structuré disponible pour tous. Il faut en cela s'appuyer sur les pratiques et les questions des DU. La demande doit être posée dans les principes d'action du Schéma d'Orientation.

L'InSHS estime nécessaire d'avoir un accompagnement solide au niveau du droit : c'est le rôle de la DAJ notamment. Des guides de bonnes pratiques, des recommandations (libre accès, données de la recherche) doivent être proposés, ainsi qu'une aide pour les chercheurs qui négocient avec les éditeurs (modèle de contrat d'édition pour les revues ; modèle de contrat de cession de droits avec volet libre accès). Une assistance juridique est également nécessaire pour la création d'outils numériques dans les laboratoires. L'InSHS a de grands besoins en matière de droit/anonymisation des données.

L'UMS Persée, qui ne limite pas son action aux corpus libres de droit et prend en charge des documents qui peuvent être soumis au droit d'auteur, souligne que les aspects juridiques occupent en conséquence une place centrale : connaître les ayants droits sur les documents et les droits dont ils disposent est nécessaire, mais également organiser les cessions de droit et assurer les conditions d'une diffusion en *open access*.

Pour les questions relatives à l'identité numérique, il est donc nécessaire de disposer de conseil et de directives précises pour les chercheurs : doit-on proposer de la formation ou encore proposer des outils spécifiques, utiliser des réseaux sociaux ? Le rôle de l'institution est essentiel pour forger les bases de production et de partage d'une IST numérique.

L'INSIS exprime un besoin plus basique (droit de la publication : quels sont les droits applicables ?). La DIST soutient le principe d'une action européenne d'accès à ces outils : des guidelines.

Plusieurs axes doivent être pris en compte, en relation avec le Collège de l'Édition publique (BSN7) : environnement juridique de l'activité de fourniture de document, droit d'exploitation des corpus, gestion des accès pour les ayants droits. L'INIST insiste sur la nécessité de clarifier le droit du web des données, de fixer en la matière des règles communes aux producteurs et aux utilisateurs d'IST. Le CCSD observe également la nécessité d'une clarification du droit applicable aux dépôts publics de la recherche, à la conservation des archives, aux services à valeur ajoutée publics.

### LES NORMES ET CONSEILS JURIDIQUES SUR LE DROIT DE L'IST NUMERIQUE

---

Le besoin d'une mise aux normes revêt plusieurs dimensions complémentaires :

- Disposer d'un schéma directeur du droit numérique applicable à l'IST, sur la base d'un cahier des charges rédigé en fonction des besoins des chercheurs définis dans les Instituts,
- Construire les bases d'un appel d'offres en vue d'une consultation juridique internationale couvrant les besoins de base,
- Mettre en place un service de conseil à la demande (Hotline),
- La DIST rappelle la nécessité de stabiliser le cadre public des services d'accès et de partage des résultats de la science publique. Or des services similaires sont proposés par le privé. Il peut donc y avoir conflit entre service public et liberté du commerce et de l'industrie. La stabilisation du droit de l'intervention publique est donc nécessaire pour fournir les services d'« Open Process », nécessaires à la valorisation des résultats de la recherche publique.

---

## REGLES DE SECURITE APPLICABLES AUX DONNEES ET PUBLICATIONS

Indépendamment des règles de conservation et d'archivage pérenne évoquées plus haut apparaît le besoin d'une action de sécurité applicable aux résultats d'IST, en particulier, mais sans exclusive, aux résultats sensibles ou confidentiels.

## LA DEMANDE DES INSTITUTS

---

La demande est présente dans les TGIR. SOLEIL fait état des pratiques mises en place pour le TGIR, et indique pouvoir s'associer à des actions de retour d'expérience et de partage de bonnes pratiques. Huma-Num indique également l'importance de cette préoccupation et les dispositions prises pour développer une action dans ce sens (sécurité de l'infrastructure comme sécurité des données). L'INS2I exprime la demande d'une action de tout l'organisme, pour lutter contre l'intrusion, l'aspiration... et pour mettre en place une action systématique de sécurisation du stockage, de préservation du patrimoine scientifique et intellectuel.

Pour l'INEE, la sécurité ne doit pas être conçue comme une « action IST », mais comme une condition de cette action.

## LE BESOIN D'UN SCHEMA DIRECTEUR DE LA SECURITE DE L'IST

---

Ce besoin se traduit par la nécessité de mettre en place :

- Un Schéma directeur de la sécurité des conditions d'accès, de conservation et de diffusion de l'IST,
- L'échange de bonnes pratiques
- La mutualisation des solutions retenues en relation avec la BSN.

## GOUVERNANCE ET PILOTAGE DES ACTIONS D'IST

---

Deux ordres de questions ont été abordés par la réflexion partagée dans le cadre du Schéma d'Orientation stratégique :

- La Gouvernance technique des projets d'IST du CNRS,
- L'appui de l'IST au pilotage scientifique des Instituts.

## LA GOUVERNANCE DES PROJETS D'IST AU CNRS

---

La gouvernance des projets d'IST présente des modalités très différenciées d'une communauté de recherche à l'autre et s'appuie sur un potentiel de service dont le trio « informaticien, documentaliste, chercheur » est la forme de base, avec des déclinaisons.

Les Instituts sont demandeurs d'une organisation lisible et efficace de la DIST sur l'ensemble du périmètre de l'IST. Il est avancé l'idée d'un « Comité de régulation » de l'IST à la DIST, impulsant les actions correspondant aux différents chantiers IST et mettant en réseau les acteurs.

L'INIST entend prendre toute la place qui lui revient dans le Comité de régulation de l'IST du CNRS, qui permettra de partager les projets entre la DIST et les unités IST et instaurer un *reporting* régulier sur les activités pilotées. Ce comité lui permettra de développer des relations en réseau avec les autres opérateurs. Huma-Num observe par exemple, qu'il faut formaliser la démarche de définition des objectifs définis pour une action donnée. Le Réseau RENATIS s'exprime également dans ce sens, et se déclare prêt à participer au Comité de régulation de l'IST.

- Mettre en place un « Comité de régulation de l'IST » intégré à la DIST. Ce Comité tripartite (Instituts, Unités Spécialisées IST, INIST en particulier, Réseaux de personnels de l'IST) aurait la tâche de coordonner le développement des actions IST du CNRS.

## L'APPUI DE L'IST AU PILOTAGE SCIENTIFIQUE DES INSTITUTS

---

Des indicateurs bibliométriques (répartition par institut, relations interdisciplinaires, part nationale ou régionale, collaborations internationales,...) sont produits par le SAPPs et sont largement utilisés. Toutefois, les instituts ne disposent pas toujours d'une vision complète de leur production scientifique (l'INC par exemple, le fait observer) et apprécieraient la mise à disposition d'outils de gestion bibliographique optimisés et partagés.

Dans ce sens, apparaît la demande d'une exploitation plus ciblée, de la base issue du WOS, par repérage des unités CNRS dans les affiliations. La demande en est exprimée par certains instituts (INSB par exemple). D'autres méthodes de mesure complémentaires sont souhaitées dans les cas où, dans le cadre des métriques existantes, la France n'apparaît pas à la place qui est la sienne dans le travail des communautés scientifiques correspondantes (c'est le cas de l'IN2P3 qui observe la nécessité de faire mieux ressortir la place qu'il occupe dans les publications au CERN). Dès à présent, des travaux d'analyses des publications ont été réalisés par plusieurs Instituts à partir du WOS (INEE, INP notamment) pour mieux connaître l'activité des unités ou détecter les collaborations entre unités.

L'INSB, par exemple, exprime la volonté de mieux valoriser son activité de recherche, en la différenciant des autres organismes de recherche travaillant dans le même domaine au niveau national voire européen. L'IN2P3 suggère de réfléchir à de nouveaux indicateurs bibliométriques : sa participation serait forte sur ce chantier de l'évaluation, de la qualité de l'IST, des formes différenciées de présentation des résultats IST (cahiers de manip, notamment), de partage de l'excellence.

- Mettre en place une offre de services d'analyse des publications des Instituts à des fins de gouvernance des projets scientifiques (suivi des collaborations, des thématiques de recherche),
- Partager ces outils avec les partenaires du CNRS, afin de développer les approches communes et les collaborations.

#### PRINCIPE D'ACTION 10 : INNOVER DANS LES DEMARCHES DE L'IST

##### LA DEMARCHE D'IST NUMERIQUE : QUATRIEME PILIER ?

Est-ce les données qui « font » la science, comme le suggère le mouvement pour un « Quatrième Pilier » de la science ou bien, comme l'avance la Contribution du CNRS à la Stratégie nationale de la Recherche, une science à fondements numériques complexes ? La question mérite une réflexion construite et partagée, que les travaux internes au CNRS (ISCC notamment) pourront stimuler.

- Animer, avec les partenaires du CNRS, un Colloque et une réflexion construite sur l'IST à l'heure numérique.

##### LA RECHERCHE SUR L'IST

La réflexion interne du Schéma a permis d'esquisser une première approche des démarches d'une ou plusieurs disciplines confrontées à la nécessité de définir les objets nouveaux de recherche qui, à l'amont de la production d'IST, définissent de nouvelles conditions de production de ses résultats. L'INSMI observe que la recherche sur l'hétérogénéité des objets et des données est un thème essentiel.

Une des missions de l'INIST est de donner accès à l'information scientifique et technique et à l'ensemble des productions scientifiques, ce qui suppose d'être en capacité d'exploiter ses gisements d'informations. Dans ce cadre, si la qualité et la richesse des métadonnées associées sont incontournables, l'exploitation directe de l'ensemble des connaissances présentes dans les textes primaires devient nécessaire, comme l'atteste l'attente de la communauté vis-à-vis des services ajoutés du projet ISTEEX. L'INIST est par ailleurs impliqué dans des projets de recherche et investissement d'avenir intégrant une part importante de fouille de textes pour proposer un accès élaboré aux connaissances scientifiques décrites dans les publications que l'INIST met à disposition de la communauté.

A l'InSHS a été proposé le lancement d'un Réseau Thématique Pluridisciplinaire sur les humanités numériques ; l'idée serait de réunir régulièrement les acteurs clef des *Digital Humanities* (Cléo, Huma-Num... mais aussi de



chercheurs dans les laboratoires) pour chercher à analyser un peu mieux les apports des Digital Humanities et les besoins qu'elles génèrent ; cela permettrait de créer un lien entre ces acteurs clef et l'Institut.

- La recherche en IST autour du Pôle Lorrain (Université de Lorraine, LORIA, ATILF...) et de l'INIST viendra en appui aux projets IST du CNRS et de tout l'ESR. Les projets retenus viseront une valorisation économique et sociale.
- Un partenariat sur l'IST sera développé par mise en réseau des projets des organismes de recherche.
- Les travaux d'amont visant l'interdisciplinarité seront développés en partenariat avec la Mission Interdisciplinarité (MI).

---

## L'INNOVATION SUR LES OUTILS ET DEMARCHES DE L'IST

L'innovation repose surtout sur l'évolution des systèmes d'information, des bases de données, des systèmes d'archivage, de la GED, des traitements linguistiques, de la conceptualisation de l'information, de la connaissance i.e. tout ce qui peut enrichir les bases de données et accélérer la recherche en ligne. Il faut être en mesure de s'emparer des technologies novatrices pour optimiser le traitement ou l'accès à l'information

L'INC fait observer que si on connaît les publications des chercheurs, c'est qu'elles sont dans des revues référencées (« Nature » etc...). Or il n'y a pas de mécanismes/d'outils/de système qui fasse paraître des publications qui n'auraient pas été référencées dans les grands systèmes. La plupart des chercheurs sont soucieux de leur visibilité et font la démarche de rassembler leurs publications (researcherID et google) : cela fait partie des choses qu'il serait bien de valoriser. Research Gate et Orcid sont en train de prendre de plus en plus leur place. A l'INRIA, des équipes travaillent sur l'ingénierie de la connaissance. Au-delà de cette question, qui relève de la valorisation des publications, l'INC a développé une démarche d'exploration sur des corpus de connaissances relatifs à un ou plusieurs objets (molécules, réactions...) et s'intéresse aux outils IST d'exploration de nouveaux champs de données et publications. Des disciplines travaillent aussi sur l'interopérabilité des données, un enjeu pour la mise en réseau et l'intégration des données en ligne qui nécessite des accords au niveau international – on peut citer par exemple l'observatoire virtuel astronomique, et le travail engagé par la Research Data Alliance, qui vient d'être formée par la Commission Européenne, la NSF et l'Australie

Le CCSD estime que, au-delà des services rendus en tant qu'outil de communication scientifique directe, l'un des principaux enjeux lié à l'innovation est de faire de HAL un outil central de l'intégration des données. Pour les usagers, il s'agit de pouvoir relier à une publication, l'ensemble des matériaux de recherche. Pour les institutions partenaires, HAL doit devenir l'outil permettant l'interopérabilité des bases de données et des référentiels utiles aux différents aspects de la gestion de la recherche.

La nécessité de disposer de nouvelles catégories d'outils d'analyse de l'IST est donc nettement présente, avec notamment la demande d'outils comportant des champs de requête beaucoup plus ouverts aux catégories sémantiques « intelligentes » (analyse des préférences, des opinions.... sur l'IST d'un champ de recherche). Cette demande est inégalement exprimée : une bonne part des communautés de la physique sont de taille et d'homogénéité d'IST suffisamment fortes pour ne pas développer un besoin de repérage sur l'IST en construction, et disposent avec Research Gate des outils qui semblent leur suffire. Cette communauté estime toutefois qu'il serait intéressant d'outiller Hal et arXiv pour extraire des articles à partir de mots clés fournis ou de pouvoir souscrire à des thèmes ou des sous-thèmes, bref de doubler l'Open Access par de l'« Open Process ».

Sans exception, les communautés scientifiques ayant contribué à la réflexion sur l'avenir de l'IST sont ouvertes à une démarche globale et demandent que les outils existants soient mieux connus et partagés, que les outils émergents soient accessibles pour tous les chercheurs: cette question, qui pose celle de l'accès à l'innovation de pointe en matière d'IST, rejoint celle de la prospective et de l'évaluation qu'aborde le premier des principes d'action posés par le Schéma d'Orientation stratégique de l'IST du CNRS.

- Faire l'inventaire et la typologie des outils d'analyse de l'IST,
- Faire connaître et partager les bonnes démarches,
- Repérer les évolutions à moyen terme et préparer leur diffusion.





Les collaborations :

BBEES	Cécile	Callou
CCSD	Christine	Berthaud
CCSD	Laurent	Capelli
Cléo	Marin	Dacos
Cléo	Pierre	Mounier
Comité d'éthique	Michèle	Leduc
DASTR	Jean-Noël	Verpeaux
DASTR	Anne-Laure	Derepas
DASTR	Anne	Sigogneau
DSI	Jean-Marc	Voltini
Huma-num	Stephane	Pouyllau
Huma-num	Marc	Renneville
IN2P3	Jacques	Martino
IN2P3	Stephane	Plaszczynski
IN2P3	Laurent	Serin
INC	Dominique	Massiot
INC	Brigitte	Diers
INEE	Stéphanie	Thiébaud
INEE	Bruno	David
Inist	Raymond	Bérard
Inist	Anne	Ciolek-Figiel
Inist	Alain	Collignon
Inist	Françoise	Drouard
Inist	Étienne	Fleuret
Inist	Claire	François
Inist	Sylvie	Gresillaud
Inist	Marie-Christine	Jacquemot-Perbal
Inist	Fabienne	Kettani
Inist	Christine	Louala
Inist	Catherine	Morel-Pair
Inist	Catherine	Poupon-Czys
Inist	Ivana	Roche
Inist	Laurent	Schmitt
INP	Jean-françois	Pinton
INP	Bart	Van Tiggelen
INP	Pablo	Jensen
INS2I	Raja	Chatila
INS2I	Michel	Bidoit
INSB	Catherine	Jessus
INSB	Ken	Takeda
INSHS	Patrice	Bourdelaïs
INSHS	François-Joseph	Ruggiu
INSHS	Odile	Contat

INSHS	Michèle	Dassa
INSIS	Jean-Yves	Marzin
INSIS	Joanna	Janik
INSMI	Frederic	Helein
INSMI	Christoph	Sorger
INSU	Françoise	Genova
INSU	Pascale	Talour
INSU	Williams	Exbrayat
Mathdoc	Thierry	Bouche
MEDICI	Anne-Solweig	Gremillet
MEDICI	Caroline	Martin
MEDICI	Clotilde	Roussel
MEDICI	Micaël	Allainguillaume
MEDICI	Olivier	Vallade
MEDICI	Thierry	Daunois
MEDICI	Virginia	Frey
Persée	Nathalie	Fargier
Persée	Viviane	Boulétreau
Persée	Eric	Astier
Renatis	Joanna	Janik
Renatis	Caroline	Abela
Renatis	Caroline	Bauer
Renatis	Bernadette	Bergeret
Renatis	Christine	Cazenave
Renatis	Marie	Chebanche
Renatis	Rodolphe	Defolle
Renatis	Gilles	Dubus
Renatis	Philippe	Eyraud
Renatis	Williams	Exbrayat
Renatis	Hélène	Falavard
Renatis	Françoise	Girard
Renatis	Christine	Hadrossek
Renatis	Bénédicte	Kuntziger
Renatis	Agnès	Magron
Renatis	Colette	Orange
Renatis	Ariane	Rolland
Renatis	Angèle	Sene
Renatis	Jean-Louis	Thomin
RNBM	Francesca	Reinardi
SGCN	Marie Christine	Balastie
Soleil	Jean	Daillant
Université de Lorraine	Jean-Marie	Pierrel