



DYNAMIQUES DE L'ÉDITION SCIENTIFIQUE, DE L'INDUSTRIE, DE L'INFORMATION, DE LA DOCUMENTATION

ACTES DU COLLOQUE
Meudon 4 et 5 novembre 2014

Un agenda 2015 pour la science publique ouverte



Direction de l'information scientifique et technique

ACTES DU COLLOQUE

Dynamiques de l'Édition scientifique,
de l'Industrie de l'Information, de la Documentation
Un Agenda 2015 pour la science publique ouverte
Meudon - 4 et 5 Novembre 2014

CONTEXTE

Ce Colloque est né à la demande des participants de la rencontre : « Innovation et gouvernance de l'IST » (CNRS, Meudon, les 18 et 19 mars 2014). Cette précédente rencontre avait conduit une centaine de représentants de tout l'ESR au constat d'une « hybridation » des usages accompagnant la publication scientifique numérique, **hybridation** qui atteint aussi bien les **résultats publiés** (données, articles-ouvrages, analyses) que les **supports** de publications et les **démarches de partage des résultats** (revues aux formes traditionnelles et nouvelles, plateformes opérant sous divers statuts...), dans un contexte où la **pluralité des opérateurs** va croissant (éditeurs nationaux, internationaux, académiques, industriels de l'information, éditeurs et opérateurs de l'Open Access...).

Le besoin d'une **orchestration collective** des changements en cours avait alors été exprimée : en réponse, ce colloque propose la mise en commun d'un « **Agenda 2015** » pour la science en phase avec le riche Agenda National sur le numérique et les plateformes.

Renaud FABRE

Directeur

Direction de l'information scientifique et technique du CNRS

PROGRAMME

Mardi 4 novembre

9h	Accueil des participants – Café de bienvenue
9h30	Présentation des journées Renaud FABRE DIST, CNRS
9h45	Ouverture du Colloque Philippe BAPTISTE, Directeur Général Délégué pour la Science (DGDS), CNRS Benoit THIEULIN, Président du Conseil National du Numérique
10h15 – 11h30	Débat - Enjeux d'avenir pour la publication scientifique Animé par Philippe BAPTISTE, DGDS, CNRS et Jean-Pierre FINANCE, CPU, Président de COUPERIN Avec : Jean-François BALAUDE, CPU, Président de l'Université Paris X Nanterre Raymond BERARD, Inist, CNRS Charles HUOT, Président du Groupement français des industriels de l'information (GFII) Claude KIRCHNER, Délégué général à la recherche et au transfert pour l'innovation, Inria Christophe SORGER, Directeur de l'INSMI, CNRS Didier TORNAY, Directeur adjoint InSHS, CNRS
11h45	Présentation des Ateliers Renaud FABRE, DIST, CNRS <i>Les partenariats - à proposer ou à renforcer en 2015 - sont définis selon trois thèmes transversaux. Une mise en commun en fin de journée permettra de dégager les bases d'un Agenda de travail commun en 2015</i>
12h – 14h	Déjeuner au restaurant « Les Erables » <i>15 bis, rue Marcel Allegot- 92190 Meudon</i>
14h – 16h30	Ateliers de travail 1 à 3
17h – 18h	Compte-rendu des Ateliers de travail Serge BAUIN – Stéphane POUYLLAU – Bruno DAVID
18h - 19H	Droit numérique et plateformes de science : Avancées 2014, projets 2015 Alain BENSOUSSAN, Avocat à la Cour
19h30	Dîner au restaurant « Les Erables » <i>15 bis, rue Marcel Allegot- 92190 Meudon</i>

Ateliers du mardi 4 novembre, 14h

Atelier 1 : Quelles exigences communes pour renforcer l'Édition scientifique ? Quels partenariats pour l'Open Access et pour l'Open Process ?

Responsables d'Atelier : Serge Bauin (DIST, CNRS) accompagné de Marin Dacos (CLEO, CNRS)

Intervenants : Dominique Babini (CLACSO), Silke Hartmann (EGU/Copernicus), Gemma Hersh (Elsevier), Hélène Lowinger (Episciences, Inria), Marc Minon (Cairn), Stéphane Plaszczyński (IN2P3, Scoap3), Iratxe Puebla (PLoS One), Jean-Marc Quilbé (EDP Sciences, EDP Open)

- A- Quelles directions de travail pour l'Édition scientifique : les choix de la BSN ?
- B- Quelles dynamiques dans les choix d'Open Access ?
- C- Quels services développer dans l'Open Process ?

Atelier 2 : Quels partenariats pour la construction des Plateformes Multi-Usages ?

Responsable d'Atelier : Stéphane Pouyllau (Huma-Num, CNRS)

Intervenants : Patrice Bellot (Open Edition Lab), Jean-Baptiste de Vathaire (Cairn), Olivier Dumon (Elsevier), Francis Lemaitre (AAR), Yvan Stroppa (Run&Share)

- A- Quelles expériences innovantes dans la construction des Plateformes Multi-usages ?
- B- Quelles politiques d'usages et d'accompagnement des usagers sur les plateformes de science ?
- C- Quels principes de régulation et Quelle gouvernance sur les Plateformes Multi-Usages ?

Atelier 3 : Quelles dynamiques pour valoriser l'information et la publication scientifique ?

Responsables d'Atelier : Bruno David (CNRS, INEE) accompagné de Francis André (DIST, CNRS)

Intervenants : Lisa Colledge (Elsevier, director of research metrics), Annie Coret (DIST, CNRS), Olivier Dumon (Elsevier), Odile Hologne (Inra), Marie-Christine Jacquemot (Inist, CNRS), Hylke Koers (Elsevier, head of content innovation), Nicolas Noquet (Thomson Reuters), Guillaume Rivalle (Thomson Reuters), Alain Zasadzinski (Inist, CNRS)

- A- Quels principes de construction des systèmes d'information sur les publications ?
- B- Quelles chaînes de valeurs afficher dans la valorisation des informations sur la publication ?
- C- Quels principes de travail afficher dans les processus d'exploration des données et publications ? Des principes de travail pour le TDM, pour la valorisation des Contenus Générés par les Utilisateurs (CGU), pour les autres formes de valorisation.

Mercredi 5 novembre

9h	Accueil des participants – café
Un grand projet en 2015 : les données publiques et la science	
9h30 – 10h	Henri VERDIER, Services du Premier ministre, Administrateur Général des données
L'information scientifique en 2015 : les facteurs de changement	
10h – 10h30	Les besoins : résultats de l'enquête auprès des Directeurs d'Unités de Recherche Francis ANDRE, chargé de mission Données de la Recherche, DIST, CNRS
10h30 – 11h	L'innovation éditoriale Elsevier (Olivier DUMON, Directeur général pour la recherche universitaire et gouvernementale), Thomson Reuters (Mondane MARCHAND Regional Sales Manager EMEA et Nicolas NOQUET, IT Account Manager), Episciences-HAL (Claude KIRCHNER, Délégué général à la recherche et au transfert pour l'innovation, Inria)
11h – 11h30	Les plateformes multi-usages : Pratiques actuelles HAL (Christine BERTHAUD, CCSD, CNRS), les projets du CNRS (Renaud FABRE, DIST, CNRS), les plateformes et la Science : grands projets émergents (Anne CAMBON-THOMSEN, Directeur de recherche CNRS)
L'information scientifique en 2015 : vers un Agenda partagé	
11h30 – 12h30	Débat Animé par Jean-François BALAUDE, Président de l'Université Paris-Ouest – Nanterre-La Défense et Didier TORNAY, Directeur adjoint scientifique InSHS Avec : Ghislaine CHARTRON, Directrice de l'INTD, GFII, CNAM Grégory COLCANAP, Coordinateur du Consortium COUPERIN François GEZE, Président du groupe des Editeurs en SHS du syndical national de l'édition Emmanuelle JANNES-OBBER, Présidente d'Eprist, IRSTEA Joanna JANIK, DIST, CNRS Christophe PERALES, Président de l'ADBU
12h30 – 14h	Déjeuner
14h – 15h	Les Acteurs de l'Édition Scientifique : bases pour un Agenda commun 2015 Alain ABECASSIS, Chef du service de la coordination stratégique et des territoires, MENESR
15h - 15h30	Allocution de clôture Renaud FABRE, DIST, CNRS

SOMMAIRE

Mardi 4 novembre 2014

Ouverture du colloque	p 06
Débat : enjeux d'avenir pour la publication scientifique	p 13

Ateliers

Atelier 1 : Quelles exigences communes pour renforcer l'Édition scientifique ? Quels partenariats pour l'Open Access et pour l'Open Process ?	p 23
Atelier 2 : Quels partenariats pour la construction des Plateformes Multi-Usages ?	p 42
Atelier 3 : Quelles dynamiques pour valoriser l'information et la publication scientifique ?	p 60
Conférence : droit numérique et plateformes de science : avancées 2014, projets 2015	p 70

Mercredi 5 novembre 2014

Un grand projet en 2015 : les données publiques et la science

Intervention de Henri Verdier, Administrateur général de données	p 72
--	------

L'information scientifique en 2015 : les facteurs de changement

Les besoins, résultats des enquêtes IST	p 77
L'innovation éditoriale	p 78
Les plateformes multi-usages : pratiques actuelles	p 81

L'information scientifique en 2015 : vers un Agenda partagé

Débat : L'information scientifique en 2015 : vers un Agenda partagé	p 86
Les acteurs de l'édition scientifique : bases pour un agenda partagé	p 93

Conclusions générales du colloque	p 97
--	------

Annexes :

Présentation d'ezPAARSE	p 98
Liste des participants	p 99

PREMIERE JOURNEE : 4 novembre 2014

Ouverture du colloque

Philippe Baptiste, Directeur général délégué pour la science, CNRS
Benoît Thieulin, Président du Conseil national du numérique

Philippe Baptiste Discours d'ouverture du colloque

Bonjour à toutes et à tous,

En tant que Directeur général délégué à la science du CNRS, j'ai le plaisir d'ouvrir, avec Benoit Thieulin, Président du Conseil National du Numérique, ce Colloque qui rassemble des représentants de tous les acteurs nationaux concernés par les :

« Dynamiques de l'Édition scientifique, de l'Industrie de l'Information, de la Documentation ».

Dans ce partenariat fort avec le Conseil National du Numérique, j'ai voulu symboliser la communauté de projets qui unit toute la recherche aujourd'hui, avec l'Institution nationale que vous dirigez, et que le Président de la République a chargé de clarifier et d'accompagner les grands changements qu'ouvre l'ère numérique pour notre pays.

Ces changements, la recherche française a l'ambition non seulement d'en tirer parti, mais de les vivre pleinement et même de les préfigurer dans certaines des directions où la société pourrait tirer bénéfice des pratiques nées dans la recherche, et dans les valeurs partagées du service public.

C'est cet ensemble de constats qui animent la stratégie dans laquelle le CNRS est moteur avec toute la recherche, et ce sont ces mêmes constats qui vous ont conduit, cher Benoit Thieulin, à accepter mon invitation pour venir voir comment l'information scientifique et technique et la politique numérique nationale peuvent mutuellement se nourrir, dans la démarche passionnante

dont vous avez la charge. Nous le souhaitons l'un et l'autre, **l'Agenda 2015** que ce Colloque se propose d'initier, sera l'occasion de mailler nos actions: ensemble, nous serons donc attentifs aux propositions qu'émettra le "Groupe d'Experts" qui sera mis en place demain soir entre tous les acteurs de l'IST.

Aujourd'hui, l'IST est à la croisée de nombreuses directions de développement qui, toutes, passent par le numérique. Parmi toutes ces directions de développement, permettez-moi de retenir trois d'entre elles qui se conditionnent mutuellement et sur lesquelles le CNRS est fier de s'engager, avec toute la recherche, avec nos amis des Organismes et des Universités.

1. La première direction est celle d'une stratégie d'effort commun de réflexion et d'action de toute la recherche; c'est un travail encouru indispensable: j'en évoquerai brièvement les valeurs et les partenariats;
2. La seconde direction, qui donne sa justification finale aux choix de stratégie, est celle qui nous mène vers une science publique ouverte aux besoins de tous ses usagers et bénéficiaires: l'IST doit traduire cet impératif global, aux formes et aux progressions diverses comme le sont les attentes des disciplines;
3. La troisième direction, celle d'une coopération plus étroite entre tous les acteurs de l'IST, est la raison d'être immédiate de notre Colloque, et ceci sous les conditions que posent les deux directions de travail précédentes.

J'accueille donc ici avec gratitude et confiance, tous ceux d'entre vous, Editeurs, Industriels de l'Information, Professionnels du numérique qui vont travailler durant deux jours et qui acceptent de relever le défi d'un Agenda 2015 commun avec les communautés de la recherche !

L'IST et l'impératif global de science publique ouverte

En lançant, le 5 décembre 2013, la stratégie IST "Mieux partager les connaissances", en invitant d'emblée la CPU et les Organismes, l'Abes, Couperin et l'ADBU à y participer dans le cadre ouvert par le MESR à travers la BSN, Alain FUCHS, Président du CNRS déclarait : "L'IST est un vecteur puissant du travail de la science : pour approvisionner les chercheurs en données et publications, pour assurer l'analyse, la diffusion et le partage des résultats, la qualité de l'IST est toujours décisive. En optimisant les échanges à tous les niveaux de notre système de recherche, l'IST peut ainsi conforter la mixité d'approches et de pratiques, qui en sont les principales richesses."

C'est cette mutualisation permanente et profonde d'une IST numérique de qualité, qui a guidé les choix de quatre Plans d'Actions Partagés (PAP) qui ont permis à cette stratégie de déboucher sur des actions et sur un engagement de toutes les communautés de recherche, en visant à "se documenter", "publier", analyser l'information scientifique", et "accompagner l'IST".

Cet effort se concrétise par le partage d'une Maison Commune nationale de l'IST qu'est l'INIST: ma toute première décision de DGDS a été d'en approuver le Projet "Ingénierie des connaissances" et je salue à distance avec chaleur et amitié son Directeur Raymond Bérard qui est absent aujourd'hui contre sa volonté: je demande aux personnels ici présents de l'INIST de l'assurer de ma gratitude et de mon admiration pour la valeur exemplaire de son action.

Dans le sens des actions entreprises sous ma direction, à l'initiative de la DIST, les Enquêtes achevées cet été auprès de 1024 Directeurs d'Unités sur près de cent questions, auprès de tous les personnels d'IST du CNRS, permettent de préciser les besoins d'IST sur un très large périmètre, que le CNRS partage avec les Universités et Organismes sur plus de 800 UMR. Sur ces bases, une étude technico budgétaire de grande ampleur sur appel d'offres européen, sera achevée au printemps 2015: cette expertise externe permettra de situer et de valoriser les besoins, d'associer les ressources entre et dans les communautés de la recherche. Avec le collège de direction du CNRS et ses dix Instituts, avec la CPU, notre stratégie commune prend corps dans les directions de travail débattues par la Table Ronde que j'animerai dans un instant, ainsi que dans les séances plénières de demain débouchant sur les bases d'un Agenda 2015. Ces choix seront également discutés lors des Ateliers de cet après-midi : je remercie vivement Serge Bauin, Stéphane Pouyllau, Bruno David, d'apporter leur expertise éminente pour animer ces regards croisés de tous les acteurs de l'IST.

Une IST servant l'objectif d'une science publique ouverte aux besoins de tous ses usagers et bénéficiaires

L'Académie des sciences vient de le rappeler dans son rapport, diffusé le 25 octobre sur "les nouveaux enjeux de l'Édition scientifique" la marche vers le libre accès aux résultats de la recherche est irréversible et indispensable, ce qui nécessite, selon la belle expression de Jean François Bach, Secrétaire perpétuel de l'Académie des sciences, dans une récente interview au "Monde", une "union sacrée internationale". N'en doutons pas, le chemin à parcourir nécessite d'avancer au contact de tous les acteurs, de tous les métiers, de toutes de les expertises sans lesquelles l'IST ne peut exister. Dans ce sens, fidèle au partenariat engagé l'an passé avec l'Académie des sciences, Renaud Fabre m'a remis le riche rapport qu'il a demandé à Serge Bauin, Chargé de Mission du CNRS pour le Libre Accès, sur les conditions de développement à retenir pour faire de HAL la véritable infrastructure nationale d'Open Access dont notre pays a le besoin.

La "Mission nationale" que l'article 2 du Statut du CNRS lui confie en matière d'IST exige que l'organisme soit aux avants postes de ces évolutions, que celles-ci soient anticipées et évaluées dans le cadre stratégique précédemment évoqué, en cherchant constamment à élargir l'éventail des bénéficiaires et usagers des résultats de la recherche publique. En assumant ces défis et cette ambition, j'ai le plaisir d'accueillir également dans ce Colloque Henri Verdier, Administrateur général des données publiques, qui partagera avec nous les visées de sa toute

récente Mission, définie le 18 septembre 2014 par le Premier ministre. Avec lui, nous avons l'ambition de réfléchir à quelques expérimentations emblématiques des formes que pourrait prendre la diffusion de données publiques produites dans la recherche publique.

Une science publique ouverte, c'est également, trouver aux plateformes de science qui se multiplient actuellement, un cadre de travail protégeant et équilibrant leurs vocations multiples. Véritable livre ouvert de la science en train de se faire, les plateformes de science intègrent données, publications, analyses et sont ainsi le carrefour de toutes les valorisations. Avec ses partenaires que sont le CERN et la DG CONNECT à Bruxelles, le CNRS entend faire le point des pratiques et s'ouvrir aux voies de régulations et de partage les mieux adaptées au devenir des plateformes de science: un prochain séminaire se tiendra dans ce sens, suivi d'autres aux Etats Unis début 2015. Maître Alain Bensoussan, que je remercie de son engagement et de celui de son Cabinet au service du projet Istex, nous apportera dans ce sens sa haute expertise sur ces sujets où nous partagerons nos approches avec le Conseil National du Numérique. Ensemble, nous recueillerons ses recommandations juridiques pour l'avenir du numérique dans la science.

Une coopération plus étroite entre tous les acteurs de l'IST numérique

Ces grands défis, parmi bien d'autres sont partagés par les acteurs de l'IST ici présents, pour beaucoup membres du Groupement Français des Industriels de l'Information, (GFII) : j'ai plaisir à saluer ici ceux des membres du GFII ici présents, ainsi que leur président, Charles Huot, directeur général délégué de Témis, qui nous rejoindra pour la Table Ronde. Le GFII mène par ailleurs, à travers ses groupes de réflexion, une action précieuse de mise en réseau, de veille et de réflexion avec les Editeurs, dans laquelle se détache la figure de Ghislaine Charton, Professeure au Cnam, que je remercie de son engagement permanent pour penser et construire l'avenir.

Dans le champ de l'IST, de nombreuses clarifications sont attendues dans les jeux d'acteurs. La gestion des contenus générés par les utilisateurs (par exemple, parcours et traces des chercheurs dans la consultation des plateformes) est traitée comme relevant encore d'une sorte de ressource "minière", au statut incertain. Par ailleurs, les stratégies d'acteurs de l'IST sont encore peu lisibles et partagées, et conduisent à de nombreuses et persistantes critiques. La communauté d'information de l'IST est pourtant très réduite et structurée dans ses standards d'information, au regard des enjeux habituels des majors: l'IST, c'est annuellement 2 millions de publiants, et pourvus de liens identifiables avec les besoins de valorisation économique et sociale.

Pourtant, l'évaluation des publications pose toujours de redoutables questions. La cohérence entre la quantité et la qualité de l'information bibliométrique est discutable, dans un monde où par ailleurs le nombre d'articles et de revues a doublé au cours des quinze dernières années !

Hétérogènes dans leur construction, leur corpus de références, leurs outils de mesures, les grands systèmes bibliométriques le sont fatalement dans leurs grilles de résultats. La qualité de la mesure bibliométrique de la science est elle-même altérée par divers biais intervenant tout au long de la filière de construction des indicateurs: **il n'y a toujours pas de standards internationaux de la mesure bibliographique**. Le CNRS, premier publiant mondial, appelle de ses vœux la construction de tels indicateurs, qui puisse un jour unifier le corpus de toutes les productions, dans une logique où l'accès à toute l'information soit universellement garanti à tous les chercheurs.

Dans le champ de l'IST en France, l'écosystème des acteurs est complexe. L'Édition scientifique nationale est représentée ici par ses grands groupements: par exemple par EDP Science, par des groupements disciplinaires dont CAIRN est une grande figure pour les SHS, mais aussi par de grands projets internes au CNRS comme l'excellent Revue.Org du CLEO. Sont également présents les grands groupes d'Éditeurs scientifiques (Reed Elsevier : à lui seul, 70% des ressources STM), les grands Producteurs d'Information (Thomson-Reuters, qui s'est d'abord développé dans la finance et l'industrie). Dans tous les cas, les Industries de l'Information sont positionnées aux points clefs de captation de l'IST et à l'interface des besoins économiques, sociaux, industriels.

Comment organiser des synergies positives entre acteurs, comment améliorer l'efficacité des collaborations éditoriales, comment travailler mieux ensemble sur les plateformes? Les questions sont très nombreuses et il faudra à coup sûr plus d'un simple Colloque pour progresser dans une meilleure intégration des projets de tous les acteurs. Puisse en tout cas cette première rencontre des professions de l'IST encourager plus fréquemment à des rencontres nationales de ce type.

Demain soir, vous mettrez en place le Groupe national d'experts auquel vous confierez le soin de rédiger les propositions d'un Agenda commun 2015 : le CNRS suivra avec attention les propositions que vous émettrez. Je forme le vœu que ces propositions nous aident à mieux nous connaître, pour mieux travailler au service de la Science publique et de l'IST. Je vous remercie par avance de cet effort collectif, que le CNRS est heureux d'accueillir.

Benoît Thieulin

Les objectifs du Conseil national du numérique (CNNum) et son calendrier

Le Conseil du numérique (CNNum) a été créé sous la présidence de Nicolas Sarkozy suite aux débats autour de la loi Hadopi rattachés à des problématiques ultra-technologiques qui ont manqué d'un organe spécialisé dans le numérique afin d'accompagner les débats. Pour pouvoir faire face à ces problématiques, une hybridation entre connaissances juridiques, connaissances techniques et connaissances des enjeux économiques était nécessaire.

Nous ne sommes qu'au début d'une révolution majeure, à l'aube de ce que Jérémy Rifkin, essayiste américain et spécialiste de prospective, appelle la 3ème révolution industrielle. Au CNNum, la composition a été élargie aux chercheurs, au monde universitaire, à des start-ups... Parmi les 30 membres, nous retrouvons des chercheurs tels que le philosophe Bernard Stiegler. Celui-ci affirme que cette nouvelle révolution est autant spirituelle que cognitive (article *Le blues du Net* <http://reseaux.blog.lemonde.fr/2013/09/29/blues-net-bernard-stiegler/>). La manière de pensée sera impactée par ces évolutions.

Internet a représenté, pour la révolution numérique et la révolution de la recherche scientifique, l'agrégation d'inventions et d'innovations. Internet est un moment particulier dans la révolution numérique. Il a été inventé par des professeurs (sur des budgets militaires américains), par de jeunes étudiants devenus professeurs qui voulaient tuer l'informatique centralisée pour mettre en place une informatique distribuée. Avant qu'internet ne s'adresse au grand public il a donc été porté par la communauté scientifique qui l'a mis en gestation. Le débat à l'époque était de décider si le choix se portait sur un réseau ouvert ou d'autres types de réseaux. Dans les années 1990, des choix ont été faits, notamment portés par Al GORE pour l'ouverture d'internet au grand public. Par ailleurs, des détracteurs comme le créateur du minitel pensaient internet trop léger, non sécurisé, trop ouvert donc dangereux, dépourvu d'un business model. Internet a pourtant été choisi car très léger et hyper égalitaire. Internet a été, à un moment donné, une vision de chercheurs et de professeurs qui ont poussés leur vision d'ouverture et de partage pour lui donner naissance. Il a été porté par des personnes qui avaient une vision politique. Aujourd'hui, il n'est pas envisageable de définir une politique déterministe sur les choix du numérique car rien n'est acquis, de nouveaux défis apparaissent chaque jour. Sur toutes ces questions nous sommes à la croisée des chemins, nous avons besoin de penser le numérique et de le maintenir dans ses choix initiaux, celui de l'ouverture. Les valeurs à l'origine de la création internet irriguent en grande partie les valeurs du numérique mais rien n'est acquis. Par exemple, l'apparition du « Cloud » remet en cause l'ouverture d'internet mais également la problématique de l'interopérabilité des plateformes qui s'inscrivent dans un débat technique et technologique.

Il faut politiser ces enjeux, en faire des débats tranchés publiquement et s'appuyer sur les travaux scientifiques pour les objectiver. Partir des sciences dures puis faire des analyses avec des sciences plus molles. Il faut mixer les sciences humaines avec d'autres sciences. Le CNNum est là pour faire le passeur entre les mondes, les communautés et les décideurs publics qui ont à trancher ces questions-là.

La conviction de Benoit Thieulin est que notre pays est difficile à réformer. Ses institutions, ses élites sont parfois en retard. C'est à la société française de peser d'avantages sur ses décideurs publics et privés. Les élites doivent être en phase avec la société. Le CNNum a par ailleurs fait une demande auprès du Premier Ministre pour le lancement d'une grande concertation sur les questions du numérique.

Les grands chantiers lancés au niveau européen (2000 – Stratégie de Lisbonne) ont été un échec total avec la dépendance de l'Europe aux plateformes américaines alors que la Russie et la Chine ont utilisé leurs leviers politiques pour s'en protéger. Par exemple en Chine avec la création du moteur de recherche WEIBO pour supplanter Google. Les plateformes chinoises protègent la Chine du monopole des Etats-Unis. Au Japon, en Corée, partout en Asie des sites, des réseaux sociaux prennent le pas sur ceux américains. En Europe, la première vague d'apparition des réseaux sociaux a été perdue en raison d'un problème de stratégie.

Le CNNum a mis en œuvre une concertation sur un grand marché numérique européen unique. Mais comment unifier l'environnement réglementaire, notamment pour les chercheurs ? Par exemple avec le « Big Data » : le marché est morcelé en autant de morceaux que le nombre d'Etats composants l'Union Européenne. Cette concertation prendra fin dans 6 mois, à la fin mars. Benoit Thieulin invite les communautés de recherche à venir participer à cette concertation. Si nous souhaitons transformer profondément le pays et les politiques il faut changer les procédés, il faut que les parties prenantes prennent la parole : il faut inviter à la table ceux qui ne l'étaient pas encore, les forces émergentes.

L'enjeu majeur est la défense de l'institution scientifique : le rapport de la société à la science doit changer, la société veut des preuves, veut comprendre... Il faut ouvrir la Science, la connaissance, les informations : il faut une **Science ouverte**.

Le lien s'est distendu entre la société française et ses scientifiques : les débats de ce colloque pourraient recoudre un peu de ce lien perdu entre scientifiques et société.

Enjeux d'avenir pour la publication scientifique

Débat animé par

Philippe Baptiste, Directeur général délégué pour la science du CNRS

Jean-Pierre Finance, Président du consortium Couperin

Avec

Jean-François Balaudé, Président de l'Université Paris X Nanterre

Raymond Bérard, Directeur de l'Inist, CNRS

Charles Huot, Président du groupement français des industriels de l'information (GFII)

Claude Kirchner, délégué général à la recherche et au transfert pour l'innovation, Inria

Christophe Sorger, directeur de l'INSMI, CNRS

Didier Torny, Directeur scientifique adjoint InSHS, CNRS

Débat

Jean-Pierre Finance ouvre le débat et indique que l'Information Scientifique et Technique (IST) connaît une révolution sous une double impulsion du numérique et de la mondialisation. La diffusion des résultats de la recherche se fait à présent à l'échelle de la planète et des tensions se développent dans ce secteur :

- Dans l'usage de la publication scientifique qui tend à diffuser la recherche avec la mise à disposition d'outils de plus en plus utilisés pour l'évaluation de l'individu, des chercheurs ;
- Dans un aspect financier : la définition de la valeur. Cette dimension a pris une place considérable avec l'envolée des prix, des coûts des revues qui sont devenus surprenants et peu maîtrisables au regard des contraintes budgétaires. En France nous sommes parvenus à pallier certaines de ces augmentations avec la création du consortium Couperin qui rassemble l'ensemble des institutions de recherche ;
- Dans l'évolution des métiers, des techniques, des utilisations : avec la modification des activités sur le terrain. La difficulté est d'anticiper les changements pour former les personnels à de nouveaux métiers, de nouvelles compétences, qui parfois modifient même l'organisation interne des établissements.

Le débat ce jour concerne :

- L'accès aux publications : libre/payant, lecture seule/Text and Data Mining (Fouille de texte) ;
- L'usage de l'IST : évaluation, pilotage, gouvernance, modification du métier de chercheur ;
- Les techniques : métadonnées, normalisation, contenus fournis par les éditeurs
- La dimension financière : qui paye quoi, la notion de propriétaire ...

Il débute donc sur la base de quelques questions majeures :

- Comment faire pour encourager les publiants à rendre accessibles leurs publications sur des plateformes ?
- Comment maîtriser les coûts ?
- Comment mettre en place le procédé de validation, le peer-reviewing ?

Selon Claude Kirchner, les chercheurs doivent communiquer les résultats de leur recherche. Tout le monde peut publier, sur ArXiv, sur des blogs. Si tout le monde publie nous nous retrouvons avec une quantité de documents qu'il faut pouvoir qualifier. Aujourd'hui on confond rendre public et qualifier. Lorsqu'on remet un document à un journal, c'est pour qu'une entité le rende public. Il est essentiel pour les chercheurs de communiquer et pour cela il faut pouvoir qualifier les publications. Ces questions se sont posées pour la plateforme HAL du CCSD qui a veillé à ce que les publiants, suite au dépôt sur HAL, puissent demander une qualification éditoriale via les Epi journaux.

Comment motiver les chercheurs à déposer ?

Pour Jean-François Balaudé il faut rendre naturelle la pratique de la mise à disposition de la publication. Il faut que les procédures soient clairement définies et œuvrer à inciter les chercheurs à déposer. La situation est différente lorsqu'on parle des laboratoires de recherche ou des universités. Quelques universités possèdent des politiques presque systématiques de dépôt. Aujourd'hui nous raisonnons à l'échelle des sites mais des universités sont également engagées dans un circuit de mise à disposition des publications. A l'Université Paris Nanterre, une plateforme sera créée en vue de permettre de déposer sur HAL ou d'aller directement sur HAL. Une réflexion est portée sur une collaboration avec le CNRS à l'échelle du site. Récemment Strasbourg a annoncé une politique systématique en la matière. Il faut sensibiliser les chercheurs aux enjeux énormes que représentent les archives ouvertes. Entamer un travail de conviction pour leur faire valoir que ces dépôts serviront leurs institutions, leur carrière. Il faut également apaiser les craintes sur la liberté des données de la recherche.

Selon Didier Torny, si nous retournons au XVIIIe siècle, à la création des revues scientifiques, celles-ci avaient deux objectifs : la diffusion et la définition des priorités. Les revues scientifiques permettaient d'identifier le chercheur à l'origine de la découverte. Ce procédé permettait une valorisation de l'ordre de propriété avec l'organisation d'une concurrence autour de la découverte : motivation fondamentale à la publication. Il faut aujourd'hui recommencer le processus et montrer l'intérêt de la publication en archive ouverte outre que celle de la révision par les pairs. Admettre que la délégation faite auprès des éditeurs n'était pas un mandat de délégation absolu. Il faut parler de la motivation des chercheurs plus que des procédés, des

contraintes autour de la publication et de l'archivage. La question de l'accès mutualisé ne peut pas être traitée par les chercheurs seuls mais avec l'appui de la puissance publique.

Claude Kirchner rappelle que pour l'Inria, le taux de dépôt des chercheurs dans des archives ouvertes, et en particulier sur HAL, était de 70-80% suite aux mesures d'incitation. Avec l'appui des puissances publiques le taux de dépôt était de 100%.

Renaud Fabre indique que comme cela vient d'être évoqué en matière des pratiques d'Open Access, les archives ouvertes offrent aux chercheurs de nombreuses possibilités comme celle d'accéder à d'autres disciplines pour féconder leurs travaux (composites, sources interdisciplinaires...).

Jean-Pierre Finance rappelle l'expérience liégeoise de Bernard Rentier : dans la promotion des chercheurs de l'Université de Liège seules les publications déposées sur l'archive institutionnelle sont prises en compte. En France peu d'établissements se sont engagés dans des processus aussi radicaux.

Transformation du métier de chercheur

Le GFII essaie de créer un espace de dialogue entre universitaires, chercheurs et industriels avec un objectif fondamental : la qualité de la publication, de la valorisation, de l'évaluation.

Selon Jean-François Balaudé le métier de chercheur se modifie et il est nécessaire d'aller vers la coercition. Pour engranger en vue de l'avenir, il faut lancer des politiques de formations des doctorants. Ces enjeux se jouent à l'échelle de la COMUE, avec la formation des étudiants, des doctorants avec la contribution éventuelle de l'INIST, des URFIST. Ces politiques, difficiles à mettre en œuvre, dépendent de la stratégie de chaque université.

Jean-Pierre Finance rappelle que pour les chercheurs, l'enjeu n'est plus seulement d'accéder à la lecture des publications mais aussi d'avoir la possibilité de travailler sur ces textes en croisant différents corpus (interdisciplinarité). Nous ne parlons plus seulement de l'évaluation des chercheurs mais de l'évolution de leur métier. Permettre le Text and Data Mining (TDM – fouille de texte) sur les acquisitions faites auprès des éditeurs fait souvent l'objet de grandes négociations. Lorsque nous achetons des revues sous forme numérique il faut savoir ce qu'on acquiert, ce qu'il sera possible de faire sur ce contenu, les droits obtenus.

Selon Claude Kirchner, la visualisation est un outil extrêmement utile pour inspirer les chercheurs qui regardent un nombre important de données. Le TDM est utile pour l'utilisation des résultats d'expérience. Des données d'autres types, extrêmement importantes, sont les données issues des évaluations des papiers, des avis. Aujourd'hui ce travail n'est pas utilisé,

n'est pas valorisé. Les rapports d'évaluation ou de lecture doivent être traités de manière encadrée et rendus disponibles. Une part de l'évaluation des chercheurs nous échappe et fini dans des bases de données utilisées lors de conférences où elles restent et sont oubliées ou dans les bases des éditeurs ou dans les bases de ceux qui les ont créées.

Didier Torny indique en exemple la base de données payante développée par l'American Chemical Society (ACS) recensant au sein du CAS Registry toutes les molécules citées dans les *Chemical Abstracts Services* (CAS - <http://www.cas.org/>) et les autres publications en chimie.

Cartographie de la recherche

La question des usages secondaires et des publications secondaires est introduite : permettre des recherches à partir des résultats de la recherche de manière secondaire ou tertiaire. Une méta-étude en épidémiologie a montré que, la plus part du temps, la recherche primaire est fautive. Nous parlons ici de Méta-validation. Quels critères restent, qu'est-ce qui dure ?

Selon Charles Huot, en IST, l'idée du Text & Data Mining rapproche le point d'interconnexion entre des domaines scientifiques qui ne sont pas forcément recouvrant avec un intérêt pour la cartographie et l'interdisciplinarité. Au-delà de la publication, la question se pose pour toutes les données de la recherche (images, données chiffrées...). Comment re-modéliser ou traiter ces données ? Ces questions ont mené le GFII à réfléchir avec les communautés à d'autres avancées possibles au-delà de la publication.

Pour Renaud Fabre le développement d'un standard de bibliographie permettant de mettre en relation les publications est essentiel : le degré d'interdisciplinarité peut être cartographié par l'identification des chercheurs cités.

Organisation

Il doit exister une cohérence dans les choix nationaux autour de l'IST et entre les différents acteurs : ABES, Inist, Couperin, la BSN, etc. Une vision collective et internationale de l'IST est essentielle. Dans le contexte numérique, où il n'y a plus d'unicité de l'objet, avoir une seule archive ouverte, une seule plateforme n'est pas envisageable, il faut envisager une multiplicité pour partager les usages.

En France se développent en parallèle HAL, les archives ouvertes de certaines universités, celles des institutions, les archives ouvertes se développent également au niveau mondial mais comment créer de la confiance dans leurs interactions ? Le dépôt multiple permet la conservation des données et commence par la standardisation des données et des

métadonnées. Il faut également que les entités de qualification travaillent en confiance avec les archives ouvertes.

Architecture

Ces problématiques récentes se posent à plusieurs niveaux : institutionnel, gouvernemental, inter-gouvernemental. Nous passons de deux décennies de pratiques, d'usages, de recherche, à un nouveau mouvement à partir des années 2000 avec la prise en main des institutions publiques. Il faut aujourd'hui construire des dispositifs techniques adéquats en s'appuyant sur toutes les expériences historiques. Comme par exemple le projet ISTEEX et sa construction sur des bases rationnelles d'accès, d'achat... Il faut que chaque micro innovation ou chaque microstructure soit immédiatement pensée dans une forme d'usage et d'interopérabilité pour la réutilisation par les communautés.

Selon le GFII, les éditeurs pourraient être parties prenantes dans ces évolutions au travers de la création et le développement d'outils. Il apparaît essentiel de tirer parti de leur expertise.

Il faut se donner les moyens de mettre en place des politiques au niveau national avec une déclinaison au niveau des sites, des institutions mais aussi suivre de nouvelles pistes qui ne pourraient pas toutes être pensées d'en haut. Par exemple avec Horizon2020 sur la gestion des données : il est possible de partir de recommandation et s'approprier un certain nombre de réflexions précises sur la gestion des données.

Périmètre

L'organisation est politique, technique mais également humaine. Nous avons réussi à mettre en place des lieux de réflexion pertinents : la BSN, des séminaires... Dans le secteur de l'IST France arrive à mixer les deux problématiques. Il faut éviter les barrières pour éviter les duplications et les concurrences inutiles. Nous ne pouvons rien faire tout seul, il est devenu donc essentiel de travailler à de nouvelles propositions au niveau européen et international.

Message de Raymond Bérard

Lors du 1^{er} séminaire de Meudon, le 19 mars, nous vous avons présenté l'état d'avancement de la réflexion sur le projet Inist « Ingénierie des connaissances ». Nous avons notamment exposé les 3 axes et 14 activités autour desquels l'Inist envisageait de structurer sa nouvelle offre de services.

1) 1^{ère} étape : validation politique du projet Inist

Depuis cette date, une étape capitale a été franchie avec la validation le 10 juin du projet par le Directoire du CNRS. Est ainsi approuvée la nouvelle stratégie de l'Inist qui s'affirme désormais comme une unité du CNRS au service de la recherche publique. C'est un tournant assez radical. En d'autres termes, nous passons d'une logique de production industrielle - qui consistait à vendre des produits documentaires - à une logique de services : l'Inist répondra désormais aux besoins des Instituts du CNRS et, plus largement, de l'enseignement supérieur et de la recherche en matière d'IST. Fortement ancrée en Lorraine avec des partenariats affirmés avec l'Université de Lorraine, l'Inist confortera ainsi son rôle national : un positionnement bien en phase avec la stratégie globale du CNRS et celle du Schéma IST « Mieux partager les connaissances », réalisé à l'initiative de la DIST et entériné au mois de décembre 2013.

Le Directoire a aussi fixé le plafond d'emploi de l'Inist : 200 ETPT (équivalents temps plein travaillés) d'ici 2016 (217 en avril 2014).

Ce projet insiste sur la notion de partenariat : travailler avec (et pas seulement pour) l'ensemble de nos partenaires, selon la logique de BSN : instituts du CNRS, autres organismes de recherche, universités. Il a d'ailleurs été reçu très favorablement par la CPU ainsi que par le Ministère.

Des décisions difficiles ont dû être prises, mais qui répondent aux exigences d'une bibliodiversité qui se nourrit de l'adaptation intelligente à l'évolution rapide des usages et des technologies :

- la production des bases Francis et Pascal s'arrêtera fin 2014.
- le Directoire a décidé de maintenir RefDoc mais uniquement pour les établissements d'enseignement supérieur et de recherche, avec la gratuité pour les unités CNRS. Le Directoire du CNRS a estimé que les chercheurs avaient besoin d'un service leur permettant de se procurer rapidement les documents difficiles d'accès (« la longue traîne »), reposant sur l'expertise de professionnels confirmés. RefDoc portera de moins en moins sur la fourniture de documents courants qui sont aujourd'hui disponibles en nombre croissant sur les portails sans oublier les achats massifs d'archives effectués dans le cadre d'Istex.

Je rejoins ici l'avis de Guy Cobolet, directeur de la Bibliothèque interuniversitaire Santé qui estime que la fourniture de documents « ne portera plus du tout sur la littérature courante, qui sera accessible en ligne, et ne concernera probablement plus qu'un noyau résiduel de fonds et de documents utiles pour l'histoire des sciences et la recherche rétrospective, qui n'auront pas encore été numérisés ».

Le secteur privé ne sera donc plus desservi. RefDoc a été suspendu le 13 décembre 2013, après l'arrêt de la Cour de cassation. Nous allons rouvrir le service dans son nouveau périmètre dans quelques jours. Il aura fallu 10 longs mois pour lever un à un tous les obstacles : juridiques, budgétaires, organisationnels, informatiques... sans oublier la concertation avec nos partenaires de l'ABES et des bibliothèques universitaires.

A travers l'arrêt de Francis et Pascal et la réduction du périmètre de RefDoc, c'est tout le modèle économique de l'Inist qui est remis en cause : ces produits étaient principalement destinés et vendus au secteur privé, industrie pharmaceutique par exemple. La stratégie de l'Inist ne sera désormais plus guidée par le chiffre d'affaire mais par les besoins IST de la communauté des chercheurs.

Mais tournons-nous plutôt vers l'avenir et les 14 activités du projet Inist. Outre certaines bien établies et qui ont vocation à perdurer (portails), plusieurs existaient déjà à l'Inist, mais demeureraient dans l'ombre telles la formation, la traduction, l'édition numérique, la numérisation, l'ingénierie documentaire, la négociation pour l'achat de ressources documentaires ou encore les études bibliométriques. D'autres sont encore émergentes comme les nouveaux usages des ressources terminologiques, l'ingénierie des langues ou la valorisation des données de la recherche.

2) 2^{ème} étape : le basculement sur le nouveau projet

Si l'étape politique a été franchie, reste une étape tout aussi délicate à franchir, celle du basculement de l'Inist sur ces nouvelles activités numériques. Elle concerne environ 75 agents (hors encadrement) dont les activités vont s'arrêter totalement ou partiellement fin 2014, et qui devront évoluer professionnellement. Un plan de redéploiement et de requalification a débuté dès juillet. Ce plan repose sur un dispositif multi-facettes que nous appelons « **plateforme RH** » :

- Une cellule RH, groupe de suivi constitué à la fois de décideurs, d'opérationnels et de représentants du personnel, qui travaille sur :
 - Les modalités RH de mise en œuvre du projet
 - le suivi de l'avancée du chantier
 - la communication

La cellule RH travaille en étroite collaboration avec le Service des Ressources Humaines de la Délégation Régionale.

- Un accompagnement méthodologique par deux cabinets de consultant : Amnyos (spécialisé RH) et Tosca Consultants (spécialisé IST) sur les volets « offre de services », « organisation », « mobilité professionnelle ».
- Des entretiens individuels menés par une conseillère RH spécialement recrutée pour l'occasion. Elle est chargée :

- d'assurer les entretiens avec les personnels, en les aidant à faire le point sur leurs compétences et à formaliser un projet professionnel,
- de les accompagner dans la recherche d'affectation en leur proposant des mesures d'accompagnement (formations, bilans de compétences).

Plus de 80 entretiens ont été menés à ce jour par la conseillère RH.

- Des actions de sensibilisation afin que les personnels puissent se positionner sur les activités qui les intéressent le plus. C'est ainsi qu'ont été organisées :
 - Une séance d'échanges avec des organismes de recherche ayant profondément réorganisé leurs services IST : Inra, Inria
 - des séances plénières de présentation des activités et des stands découverte qui ont rencontré un vif succès
 - des rencontres avec le CCSD, le Cléo et Persée qui ont présenté les pistes de partenariat avec l'Inist

- Des immersions dans des structures IST externes dont l'objectif est de favoriser la compréhension des problématiques IST dans un autre contexte, tous métiers confondus : ingénieurs documentalistes, informaticiens, administratifs. Sont prévues ou déjà réalisées des immersions à l'Université de Lorraine, à l'ABES, au CCSD, à la BnF, à la Bibliothèque Royale de Copenhague.

- Un plan de formation qui sera finalisé une fois que les nouvelles affectations seront connues.
Un budget spécifique est prévu à l'Inist et à la DR, la DIST intervenant également pour mobiliser un maximum de ressources de formation.

3) Calendrier

- Début décembre 2014 : production d'un nouvel organigramme rendu nécessaire par l'arrêt de services qui constituent l'ossature des actuels pôles métiers.
Il est acquis que cet organigramme obéira largement à une logique « projet ». Il s'appuiera sur les conclusions d'un groupe de travail « gestion de projet » mis en place en novembre 2013 pour résoudre des dysfonctionnements.
Le scénario suivant (appelé « task force ») a été retenu :
 - Mise en place pour chaque projet d'une équipe multi-métiers autonome dédiée à la réalisation du projet
 - Le chef de projet gère les ressources du projet (RH, y compris entretien annuel d'activité, et budget).

- Les membres des équipes projet (développeurs et fonctionnels) affectés au projet sont détachés de leur service d'origine pendant la durée du projet
 - Cette « task force » intégrera une dimension innovation : réalisation d'études et prototypes.
- Janvier-mars 2015 : lancement des bourses internes par vagues (encadrement supérieur suivi par l'encadrement intermédiaire puis les équipes), affectations des agents. L'organigramme fixera des fourchettes d'effectifs pour chaque activité afin de ménager la nécessaire flexibilité entre activités, notamment émergentes : les besoins RH seront en effet appelés à évoluer en fonction des projets.

4) Travail sur le volet « offre de services »

Parallèlement à cette plateforme RH, nous travaillons sur le volet « offre de services ». En effet, il est apparu que la description des activités n'était pas toujours très claire à la fois pour nos partenaires extérieurs et pour les agents en repositionnement. Les fiches activités sont le fruit d'une réflexion technique conduite d'un point de vue interne et expert qui a abouti à des formulations ne donnant pas toujours à voir le type de service proposé à la communauté scientifique.

Il a donc été jugé nécessaire d'engager un travail de :

- traduction de certaines de ces activités afin de les rendre plus compréhensibles
- et de reformulation à partir d'une matrice commune en vue de comprendre les interactions entre les différentes activités.

Le travail porte sur :

- La nature des services et prestations (« quoi ?»)
- Les cibles (« pour qui ?»)
- La volumétrie (« combien ?»)
- Les modalités (« comment ?»)

Ce travail de traduction et de reformulation est variable selon les activités :

- les cibles et volumétries sont bien connues pour des activités pratiquées de longue date comme RefDoc, les portails. Il en est de même pour la négociation de ressources électroniques qui s'exerce dans le cadre de Couperin depuis que le CNRS a adhéré au consortium.

- Elles restent à préciser pour des activités existantes mais peu connues :
 - Traduction : une équipe de 3 personnes peut-elle répondre aux sollicitations de tout l'ESR ? Tous les chercheurs peuvent-ils demander à faire réviser leurs articles ?
 - Formation : faut-il viser tous les champs de l'IST ou bien se concentrer sur les domaines d'expertise de l'Inist ?
- Activités émergentes (gestion des données de la recherche, terminologie, ingénierie des langues) : leur caractère émergent laisse augurer des montées en charge progressives, selon le niveau de maturité des projets et des orientations données par BSN.
- Activités nouvelles (secrétariat de rédaction) : l'Inist a été sollicité par une revue publiée par le Muséum qui perd un emploi de secrétaire de rédaction. Un cadrage est nécessaire à cette activité de substitution : jusqu'où aller ?

Un cadre est indispensable à toutes ces nouvelles activités. Faute de quoi, l'Inist va rapidement se trouver en défaillance vis-à-vis des demandeurs.

Une autre difficulté tient à la distinction entre ce qui relève de programmes/projets nationaux confiés à l'Inist (comme les projets de type ISTEX) et ce qui relève de l'offre de services à qualifier par l'Inist, comme les sites web, l'appui au pilotage, etc.

La gestion des données de la recherche fait actuellement partie de la 2^{ème} catégorie car initiée par l'Inist avec des laboratoires lorrains. Elle peut être appelée à glisser vers la 1^{ère} catégorie (projets nationaux) si un rôle était attribué à l'Inist par BSN dans ce domaine.

L'objectif est de formuler une offre de services clairement lisible pour la communauté scientifique et de répondre ainsi aux objectifs du plan stratégique en cohérence avec l'étude technico-budgétaire conduite par la DIST qui aidera à définir ce cadre.

Atelier 1
Serge Bauin et Marin Dacos

Quelles exigences communes pour renforcer l'Édition scientifique ?
Quels partenariats pour l'Open Access et pour l'Open Process ?

Intervenants	Appartenance
Dominique BABINI	CLACSO
Serge BAUIN	CNRS-DIST
Marin DACOS	CNRS-CLEO
Silke HARTMANN	EGU-Copernicus
Gemma HERSH	ELSEVIER
Hélène LOWINGER	EPISCIENCES-Inria
Marc MINON	CAIRN
Stéphane PLASZCZYNSKI	CNRS-IN2P3-SCOAP3
Iratxe PUEBLA	PLoS One
Jean-Marc QUILBE	EDP sciences-EDP Open

Présentation de Silke Hartmann
Innovative open access publishing and public peer review

Lors de sa fondation en 1988, Copernicus a commencé par proposer l'organisation de conférences et des services pour les chercheurs. **Copernicus est cofondateur d'OASPA [Open Access Scholarly Publishers Association].**

En 1994, Copernicus publie sa première revue et en 2001 l'ensemble de ses revues passe en open access et open review

Mode de travail de Copernicus :

L'Ouverture est un principe fondamental pour Copernicus qui opte pour la stratégie OA3, avec le

- Libre accès aux publications
- Libre accès à la gestion et au processus d'édition
- Libre accès à la révision par les pairs

Tous les services sont internes, l'ensemble des processus éditoriaux précédant la publication est important. Il y a deux étapes avant la publication. L'évaluation par les pairs étant en libre accès, la discussion est publique et toute la communauté peut y contribuer.

Il y a au moins 2 rapporteurs désignés mais la communauté peut contribuer à tout moment à la discussion. Au bout d'un certain temps la phase de discussion est close pour aboutir à la version finale. Les services offerts par Copernicus se trouvent à plusieurs niveaux : correction d'épreuves, édition/typographie, processus de révision par les pairs.

Aspects financiers

Il existe des modèles de licence pour les 36 revues de Copernicus dont 5 seulement appartiennent à Copernicus. Copernicus travaille ainsi avec d'autres communautés et laisse aux chercheurs le choix de la licence la plus appropriée. Les communautés et sociétés savantes font leurs propres choix financiers : soit faire des bénéfices, soit subvenir aux frais de l'ensemble du processus et faire ainsi en sorte que l'auteur ne participe pas du tout au coût de publication.

APC : Copernicus veille à ce que les APC demandés ne soient pas trop élevés. Le coût de l'article dépend du nombre de pages et il n'est pas demandé de frais supplémentaires pour la visualisation. Le chercheur peut publier à moindre frais en fonction de la longueur de l'article (500-800€/10 pages d'article). Les accords institutionnels de Copernicus précisent que ce n'est pas à l'auteur de s'acquitter des frais. Copernicus travaille directement avec l'institution à laquelle il appartient.

Présentation



Innovative open access publishing and public peer review

Silke Hartmann

Copernicus Publications | November 2014

@copernicus_org

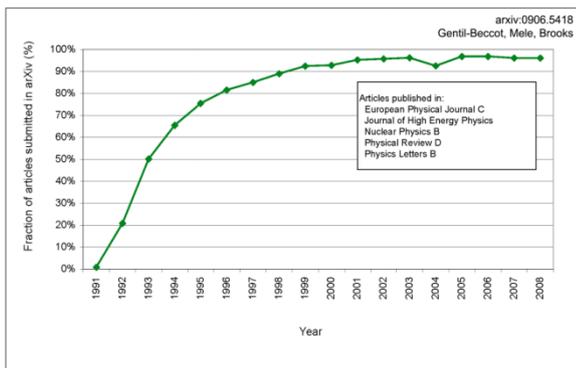


Présentation de Stéphane Plaszczyński

Publication en open access, un service prépayé de certification

SCOAP³ est une organisation issue de la communauté de la physique des particules qui organise le rôle des revues dans cette communauté. Il s'agit également de savoir si d'autres communautés peuvent s'organiser de cette façon.

La communauté de la physique des hautes énergies (HEP) est pionnière dans l'utilisation des articles, ArXiv a été créée dès 1991. 1989 voit la naissance du web au CERN et le protocole http devient le mode de communication dès les années 90.



La communauté des PHE est une société fermée, ce qui facilite les calculs de statistiques. Avec la démocratisation d'internet, on constate en 2000 une augmentation du nombre de revues

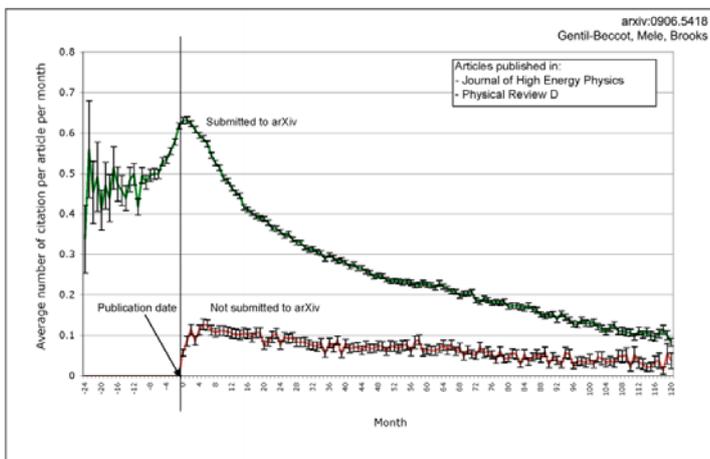


Schéma indiquant les citations des articles (ci-dessus) : La ligne verticale désigne le moment de la publication dans le journal. Pour les articles qui sont dans ArXiv, qui est le vecteur d'échange de la communauté HEP, les citations commencent avant la publication dans le journal. On constate ainsi que le preprint est un accélérateur de connaissances en informant les chercheurs dès la parution ce qui permet sa citation avant publication, avec un pic de citations à la publication.

De même, on montre qu'un article paru en preprint dans ArXiv est beaucoup plus cité qu'un article qui n'a jamais été déposé dans ArXiv.

Le lecteur d'ArXiv reçoit la liste quotidienne ou hebdomadaire de tous les nouveaux preprint publiés dans ArXiv. S'il est intéressé, il va voir les références qui renvoient à d'autres articles et

ainsi toute une toile de liens se tisse. L'outil spires du CERN est en train d'être remplacé par Inspires qui contient toutes les métadonnées de ces articles.

Créer des passerelles interdisciplinaires : Lors de ses recherches le chercheur rencontre des aspects qui lui sont méconnus. Alors, il cherche l'information en passant par des plateformes telles que google ou wiki, car il manque des passerelles interdisciplinaires.

La disparition des bibliothécaires appauvrit l'accès à l'information. En effet, l'accès immédiat à l'information sans possible errance et chemins de traverse conduira à l'appauvrissement de la connaissance. L'efficacité a un coût. En recherche les chemins de traverse ont du bon, car ne pas accéder directement à l'information la plus pertinente permet de s'égarer et d'accéder à une connaissance annexe, qui pourra rester en mémoire.

Pourquoi des journaux ? Le preprint, une fois publié évolue et plusieurs versions sont déposées dans ArXiv. Cela permet d'accéder à des articles plus complets. Mais l'inverse existe aussi : il arrive que l'article publié dans un journal après peer review ne soit pas mis à jour sur ArXiv.

Le journal est un gage de qualité mais ArXiv a un système d'auto-régulation par les pairs : il n'est pas possible d'y publier un article non robuste. Les revues sont financées par la communauté pour organiser l'évaluation. En effet, si l'évaluation est celle de la communauté d'arXiv uniquement, il pourrait y avoir des problèmes de conflits d'intérêts. Le besoin est bien celui de l'évaluation et non pas la publication.

SCOAP³ est une organisation de la communauté de Physique des particules qui a rassemblé les crédits nécessaires et finance le service d'évaluation des éditeurs.

Scoop3 a permis de se libérer des APC. L'article appartient au chercheur qui en récupère le copyright et peut être archivé. L'éditeur n'a pas le droit de le vendre et donc il réduit les abonnements d'autant. Les éditeurs ne sont pas mécontents, certains gagnent en stabilité financière. Le business plan de SCOAP³ est complètement **transparent pour le chercheur**.

C'est une communauté pour laquelle les épi-journaux auraient bien fonctionné. Mais au moment de la construction de SCOAP³, il n'y avait pas d'avantages pour un chercheur de s'investir dans le travail d'évaluation pour sa carrière ou sa reconnaissance.

SCOAP³ a été inauguré en janvier 2014 et à ce jour, on compte 3630 articles qui ont été financés mondialement. 50% de la production mondiale de cette communauté est en OA : ce qui équivaut à 50% de 7 8 000 articles. 82 pays en bénéficient. Une réduction a été accordée à plus de 2500 établissements.

Etendre SCOAP³ aux autres communautés ?

La communauté de PHE est

- petite : 7 à 8 journaux
- déjà green
- possède un organisme (international) qui soutient le projet.

De plus, en physique des particules, le nombre de publications a atteint un seuil stable depuis quelques années.

On peut noter que la communauté d'astrophysique a également atteint une asymptote pour le nombre de publications.

Présentation

Place des journaux en HEP et extension possible du modèle SCOAP3

S. Plaszczyński

SCOAP3 Network/Conseil Point chargé de mission CNRS, CORISTANbPa

Colloque Meudon, 4-5 Novembre 2014



Présentation de Hélène Lowinger

De l'archive ouverte à l'Epi-journal, entre rupture et continuité

Enjeux des revues

Elles certifient, enregistrent et labellisent. Elles déterminent la paternité des travaux de recherche. Elles sont utilisées dans la production d'indicateurs bibliométriques, et concourent à la carrière des chercheurs (recrutement, promotion, notoriété).

Enjeux financiers : leur coût est en constante augmentation d'autant que le nombre d'articles et de revues est en inflation.

Le système est-il équilibré ? On peut avoir l'impression d'un déséquilibre. En effet, du côté recherche : les chercheurs font la recherche, écrivent l'article, relisent, évaluent. De l'autre côté : est-ce que le travail d'édition est proposé au bon tarif par les éditeurs ou les sociétés savantes ?

Naissance de l'OA : en 1991 naissait ArXiv, en 2001 HAL et en 2003 PLOS.

Une épi-revue consiste à construire un service de publication scientifique via une infrastructure publique comme **Epi-sciences** construite par le CCSD et qui s'appuie sur HAL, ArXiv et CWI, éventuellement d'autres archives.

Cette organisation offre un accès libre immédiat et sans embargo aux résultats scientifiques, sans contrepartie financière. C'est une alternative à l'édition classique.

Dans la continuité de l'édition classique, l'épi-revue est dotée d'un **épi-comité** qui garantit une cohérence et un haut niveau de qualité des contenus scientifiques accessibles à tous.

Dans la rupture avec l'édition classique, l'épi-revue permet la reprise en main des revues par les communautés scientifiques, et s'adresse aussi bien à des revues nouvelles émergentes qu'à des revues qui existent déjà et souhaitent faire évoluer leur modèle éditorial.

IAM (informatics and applied mathematics) et Episciences math sont les 2 épi-comités qui existent actuellement.

C'est un outil souple et robuste. Cette plateforme permet les échanges avec des objets plus variés que ceux des revues papier similaires : des images, des codes, des vidéos, en liaison avec l'article.

Indépendance financière : Le chercheur retrouve son mot à dire, et n'est pas lié financièrement.

Le **Workflow est traditionnel** :

1. L'auteur dépose son article dans une archive ouverte associée.
2. Il soumet l'article sur la plateforme en choisissant l'épi-journal qu'il souhaite et en utilisant l'ID de l'archive.
3. Le rédacteur en chef assigne des relecteurs qui commencent leur relecture.
4. L'article est accepté ou ne l'est pas.
5. Puis a lieu un travail de mise en forme si nécessaire ainsi que de communication.

Pour aller plus loin :

D'autres revues peuvent être candidates.

La parution des articles peut se faire au fil de l'eau ou bien par numéro une fois que plusieurs articles sont révisés.

Inria travaille étroitement avec le CNRS pour que la plateforme s'adapte à tout journal qui se manifeste.

Le rédacteur en chef possède des autorisations particulières de connexion (pour gérer la revue, assigner les relecteurs, connaître l'étape, faire des relances - les relances automatiques sont possibles).

La licence est en CC-BY. L'épijournal, qui a un ISSN, permet l'attribution d'un DOI à l'article.

L'épi-journal permet plus de discussions et de partage que dans une revue classique. C'est du green open access.

L'épi-journal entraîne aussi une **réduction des délais de publications**.

Perspectives : à court terme, arrive la possibilité de revues "augmentées" qui enrichissent la publication en ajoutant des programmes ou des algorithmes (Exemple : IPOL qui met en ligne le programme informatique et le jeu de données et permet d'exécuter les programmes).



De l'archive ouverte à l'Epi-Journal, entre rupture et continuité

Hélène Lowinger, Inria

Colloque CNRS 4-5 novembre 2014 : Dynamiques de l'édition scientifique,
de l'industrie de l'information, de la documentation.

Présentation de Iratxe Puebla

Nouveau process, émergence d'un nouvel objet éditorial, la mégarevue

PloS One publie surtout en médecine. Tous les résultats doivent être rendus publics. Le processus de soumission à une revue et d'évaluation par les pairs implique une série d'aller-retours entre les pairs jusqu'à ce que l'article soit finalement accepté dans la revue choisie et peut-être aussi dans une revue qui n'est pas celle voulue en premier lieu. Cela implique beaucoup de temps pour être publié. L'innovation de PloS One réside en son modèle éditorial : le peer review est concentré sur la vérification que l'article est "**scientifiquement rigoureux**". L'éditeur et l'évaluateur connaissent l'audience spécifique à cibler et le niveau d'impact de l'article.

Points communs entre tout mégajournal

- Couvre toute discipline
- Publié quotidiennement
- Pas de limite de taille, de volume pour l'article
- On peut ajouter des objets à l'article (jeu de données, etc.)
- Licence cc-by
- Chaque article paye pour couvrir son propre coût (1350\$ pour PloS One)

PloS One a connu une croissance des publications dès sa naissance, accrue de façon très importante après l'attribution d'un Impact Factor.

Ce modèle connaît ainsi un succès durable. Depuis le lancement de PLOS one, beaucoup d'autres 'megajournal' ont été créés la plupart en sciences médicales mais pas seulement.

Particularité des mégajournals (outre leur taille) :

- Permet à toute recherche qui a des difficultés à être publiée comme des "résultats négatifs" ou qui n'est pas dans la ligne éditoriale des journaux existants, d'être publiée. Cela remet toutes les disciplines au même niveau, la recherche s'effectuant sur l'ensemble des publications, on retrouve la même visibilité pour tous les champs disciplinaires

Avantages de ce modèle éditorial :

- Faciliter et favoriser la parution des résultats et découvertes académiques par des délais de publications courts
- Modèle viable et autosuffisant.
- L'échelle importante du mégajournal permet de proposer des prix bas
- Permet de favoriser les découvertes en autorisant la navigation dans un corpus complet de ressources publiées

Le défi : comment évaluer l'impact des articles dont l'intérêt n'est pas évalué avant la publication?

Le travail est fait au niveau de l'article et non pas du journal par l'utilisation d'une nouvelle métrique qui porte sur chaque article. Il s'agit de l'article level metrics (ALMs).

Les mesures peuvent se faire sur les usages (vues, téléchargements (PDF ou XML), sur les vues sur PMC (téléchargements à partir de PMC) sur les citations (PMC, crossref, SCOPUS, WOS), sur le bookmarking (citeUlike, Connotea, Mendeley), sur l'utilisation dans des blogs, réseaux sociaux (twitter, facebook), sur les discussions (commentaires PLOS ou les blogs de google).

Présentation



MegaJournals

Disruption, innovation and the future of scientific communication

Iratxe Puebla
Deputy Editor PLOS ONE
4 November, 2014

Présentation de Jean-Marc Quilbé EDP Open/A&A lien aux données

EDP Sciences représente le bras armé des sociétés savantes STM françaises dans le monde de l'édition.

Aujourd'hui 5 sociétés savantes sont actionnaires d'EDP Sciences. Les actionnaires sont des scientifiques et participent à l'évaluation de la recherche. Le terme de société savante souligne la différence avec un éditeur privé.

EDP Sciences publie dans 3 grands domaines : journaux académiques, livres et EDP open (une plateforme en Open Access). Elle publie 12 journaux soit en modèle gold ou soit en platinum et des journaux de santé. **Cette plateforme comprend journaux, actes de conférence et livres.**

Elle est configurable pour présenter différents modes éditoriaux. Les workflows sont adaptés en fonction des besoins des différentes communautés (médecine, astrophysique ou biologie). Sur les 200 000 articles présents sur la plateforme, il est indiqué ce jour que près de 90 000 sont en libre accès, (gold ou platinum). Ils peuvent provenir de données issues de journaux classiques ayant été libérées après les périodes d'embargo. EDP Sciences s'aligne sur les embargos indiqués par la commission européenne, soit 6 mois pour les sciences STM.

EDP Sciences adapte également les licences à ce qui est demandé aujourd'hui. Les revues en open access peuvent comporter 3 types de licence -CC ou le Copyright du propriétaire de la revue (ex revue en OA de l'ONEMA, garde le copyright de l'ONEMA).

EDP Sciences prépare également le **passage des revues classiques en Open Access** et travaille avec les sociétés savantes concernées pour les assister. Plusieurs titres en OA gold sont créés par an.

EDP Sciences ne possède pas d'autres sources de financement que ceux émanant des revenus des abonnements et des APC. Ce sont les abonnements qui financent la transformation du modèle éditorial de l'entreprise et donc EDP Sciences a **intérêt à maintenir cette source de financement le plus longtemps possible.**

La revue KMAE (Knowledge & management of aquatic ecosystems) qui jusqu'en 2009 était connue sous un titre français transformé avec l'ONEMA avait un format classique avec un facteur d'impact très bas et depuis son transfert elle a évolué positivement. Il en va de même pour la revue « Parasite », dont les contenus et le titre sont désormais en anglais, et qui est financée par son institution. EDP Sciences n'apporte rien d'autre que sa prestation de service.

Depuis que les revues sont passées en Open Access, on constate un accroissement important de la visibilité et des citations.

Les possibilités de mise en place du libre accès sont variables selon les domaines. EDP Sciences coopère avec des archives ouvertes, comme ArXiv pour la physique. Tous les articles d'astrophysique et de physique sont déposés dans ArXiv. Depuis peu les articles publiés par des agents du CEA sont déposés dans HAL. EDP Sciences se situe ainsi bien dans les préoccupations des scientifiques.

Le passage en Open Access est plus facile si la revue a une certaine notoriété qui lui apporte des parrainages ou des APC (Article Processing Charges). Cela permet un premier filtre. Il faut également que la communauté d'auteurs soit prête à payer pour le service, ce qui n'est pas toujours le cas.

Il existe des revues, dites, prédatrices qui font beaucoup de démarchages auprès des chercheurs.

Cost est un exemple de revue issue d'un projet européen qu'EDP Sciences a repris à sa charge à la fin du contrat européen. Cette revue a obtenu son premier facteur d'impact (de 2,5, ce qui est excellent dans son domaine) cette année après 3 années de publication.

EDP Sciences a réalisé une étude avec un partenaire d'Oxford, qui cherchait à déterminer la place des sociétés savantes et de leurs publications dans le monde.

Les 33 sociétés questionnées ont décidé d'attendre que l'Open Access soit la norme. C'est un nouvel environnement auquel ces sociétés doivent s'adapter. L'Open Access aura un impact fort sur les revenus des sociétés savantes qui dépendent des publications.

Toute la communauté de recherche, a besoin d'un label, certifiant la qualité des publications et les sociétés savantes ont un rôle à jouer dans cette configuration <http://www.edp-open.org/> EDP open survey.

EDP Sciences publie également des livres en Open Access. EDP Sciences souhaite que les sociétés savantes soient à nouveau au cœur du processus éditorial puisqu'elles sont composées de scientifiques. Il ne s'agit pas de concurrence mais de complémentarité entre les acteurs.

Présentation



Publishing partner of the scientific communities

“EDP OPEN”

la réponse Open Access des sociétés savantes

Jean-Marc Quilbé
Président
jean-marc.quilbe@edpsciences.org

CNRS 4 novembre 2014

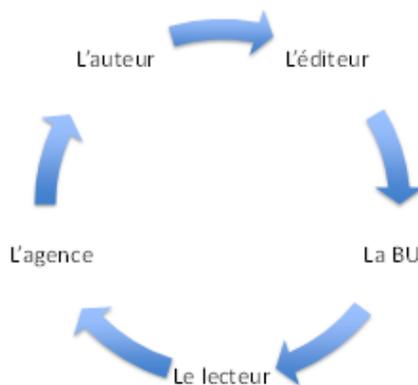
1

Présentation de Marin Dacos

De la séparation des pouvoirs. L'open access au service de quel acteur ?

Actuellement le financement de l'Open Access et les modèles qui semblent s'imposer concernent les disciplines dites de sciences dures. Il s'agit du modèle auteur payeur (APC, Article Processing Charges) qui interroge les SHS car il semble difficile à transposer dans le domaine et pose certaines questions et réflexion sur les dérives potentielles du modèle Auteur/payeur. Marin Dacos regrette d'ailleurs que les Presses universitaires ne soient pas présentes.

Qui est au centre du système ?



Est-ce l'auteur ? Est-ce l'éditeur ?

L'auteur effectue la recherche, rédige l'article. En cela, il constitue le centre du système.

L'éditeur sélectionne les articles, les diffuse, leur assure une audience et une labellisation. En cela, il constitue le centre du système.

La bibliothèque universitaire sélectionne les revues en les acquérant, les met à disposition des chercheurs et étudiants. En cela, elle constitue le centre du système.

Le lecteur lit les articles, les cite. En cela, il constitue le centre du système.

Mais il y a un autre acteur qu'on oublie souvent :

L'agent de financement, il finance la recherche, demande des livrables, qui sont souvent des articles, et donc il finance la production de ces articles. En cela, il constitue le centre du système.

Du point de vue de l'open edition, personne n'est au centre du système, c'est le système qui lui-même est fonctionnel, à partir du moment où il assure une séparation des pouvoirs.

La séparation des pouvoirs dans un monde démocratique permet d'éviter les risques des dérives d'un acteur par rapport à un autre, c'est une autre manière de dire qu'un monopole serait dangereux. Elle permet à chaque acteur de développer sa logique et si un acteur a des difficultés et oublie sa fonction et dérive, on peut imaginer qu'il y aura un rattrapage par les autres acteurs.

Dans un système auteur payeur APC on n'est plus dans ce genre de système puisque l'agence de financement finance l'auteur qui finance l'APC qui finance la revue. PloS One suscite

admiration et inquiétude mélangées car l'APC (la contribution financière) est au centre de ce système.

Le lecteur est en général oublié dans ce mécanisme, la publication relève d'une commande de l'auteur auprès de l'éditeur : le lecteur ne joue plus aucun rôle dans ce système et on assiste à un emballement économique proche. Par exemple en 2013, l'Université Aix Marseille a versé à PloS One 114 000 € de frais de publications. Cela conforte les acteurs les plus puissants à affaiblir les petits (« effet saint Matthieu ») et, probablement à construire de nouveaux trusts. A cela s'ajoute le risque que « l'autorat » devienne coûteux (sans crédits pas de publications). On sait que dans cette hypothèse, un certain nombre de laboratoires, dans les disciplines SHS, seront dans l'incapacité financière d'opérer sur ce mode.

La nécessité de payer n'est pour autant pas contestée. Le travail éditorial doit être rémunéré. **Mais comment réguler le système au mieux ?**

On pourrait assister à une double embolie : si l'on cite Elsevier le tout Gold serait une hypothèse de travail où Elsevier verrait une croissance continue de ses revues classiques payantes auxquelles l'éditeur ajouterait des articles en Open Access. Il est possible que la facture double. Il faut donc que l'on **soit attentif aux équilibres généraux et aux régulations.**

Le système Open Access idéal est celui où le **lecteur est seul au centre du système**

Dans ces conditions c'est PLOS one qui est l'un des acteurs qui pousse les article-level Metrics (ALMs). Open edition a eu des échanges avec Martin Fenner d'ALM <http://blog.martinfenner.org/2013/12/11/what-can-article-level-metrics-do-for-you/> et un travail s'est engagé avec lui. En effet les présupposés de la biologie et de la médecine ne sont pas les présupposés de toutes les disciplines. Il faut donc convaincre que dans le modèle ALM il y a des repères qui ne sont pas totalement réversibles.

Si on revient à un vieux modèle de publication : la publication produisait un usage qui produisait des ventes qui produisaient un revenu et qui alimentait un système vertueux dans lequel l'éditeur pouvait continuer à exister.

Dans le système auteur payeur qui dériverait, le financement alimenterait la publication et l'usage disparaîtrait, il ne serait plus au centre.

Il faut donc imaginer intégrer un tiers dans le système qui régule le système et assure une séparation des pouvoirs.

Il faut imaginer un mécanisme où l'usage alimente le système et non pas le besoin de publier.

Il existe 4 initiatives (unglue.it, OpenEdition freemium et Knowledge Unlatched) où l'on réintroduit un tiers qui paye (la BU ou le citoyen chercheur) permettant un système ternaire et non binaire. Ces initiatives ne sont certes pas adaptées à toutes les situations, mais elles ont le mérite d'assurer une séparation des pouvoirs.

Cette séparation des pouvoirs est une clé vers la **bibliodiversité** qui doit rester un objectif : diversité des auteurs des disciplines, des langues, des dispositifs de publications.

Il ne faudrait pas qu'au passage de l'Open Access les SHS disparaissent dans le flot.



De la séparation des pouvoirs L'open access, au service de quels acteurs?

Marin Dacos, Centre pour l'édition électronique ouverte, directeur

4 novembre 2014 - Meudon

Colloque « Dynamiques de l'édition scientifique, de l'industrie de l'information, de la documentation.

Un Agenda 2015 pour la science publique ouverte. »



Présentation de Dominique Babini

Amérique latine : une tradition de libre accès non commercial

Le réseau central de CLACSO regroupe 365 centres de recherche, dans 21 pays de la région de l'Amérique latine. Ce réseau comprend 400 revues dont 68% sont en libre accès.

Il permet d'avoir une idée de la recherche dans cette région où la recherche est principalement financée par des fonds publics (pour 2/3).

Il y a une forte tradition en Amérique Latine liée aux systèmes d'information des Nations Unies et aux systèmes coopératifs d'information. Ce sont des archives ouvertes thématiques qui chaque fois offrent en plus du texte complet.

Le réseau propose un libre accès aux revues qui n'ont pas d'APC. Dans cette région, les revues sont principalement publiées par les universités, les sociétés savantes, les organisations et les associations professionnelles. Pour le libre accès, les revues ont choisi des portails, principalement SCIELO (sciences dures majoritaires) et Redalyc (SHS majoritaires).

Sur ces 2 portails, on compte 1300 revues. La publication en libre accès est gratuite.

Ils offrent des indicateurs bibliométriques et scientométriques aux revues. Ce qui a permis d'améliorer la qualité des publications dans la région.

Concernant les portails OJS (open journal systems) on peut donner 3 exemples de portails : São Paulo, Chili et Mexico. Chacun de ces portails héberge plus de 100 revues publiées par l'Université. C'est elle qui gère le portail.

Plus récemment, ont été développées les archives ouvertes institutionnelles. Elles se sont organisées en un groupement : La « referencia », similaire à Openaireplus en Europe. Elle regroupe 9 pays et possède un moissonneur régional. C'est le début du green Open Access pour l'Amérique latine.

Le Pérou, l'Argentine et le Mexique ont déjà fait passer leur loi du libre accès vert qui exige que toute recherche financée par des fonds publics (en Amérique latine, les 2/3) soit déposée dans des archives ouvertes. Au Brésil et au Venezuela, ces lois sont en débat au Congrès.

Défis pour l'Amérique latine :

- Transformer les recommandations en obligation
- Utiliser des licences qui permettent la réutilisation
- Débattre du bien-fondé de l'APC - L'Amérique Latine peut-elle se permettre de payer des APC calculés sur des référentiels de coûts qui ne sont pas les siens, et peuvent représenter plusieurs mois de salaires ou la moitié d'une subvention de recherche

En effet, si l'Etat paye les auteurs, paye les évaluateurs des articles, peut-on développer un système de libre accès qui soit à l'image des systèmes de santé et d'éducation ? Il existe des systèmes de santé et d'éducation qui sont gratuits : est-ce qu'on pourrait avoir un système de libre accès qui soit gratuit pour les 2 fonctions de **lecture** et de **publication**. Puis au-dessus de cette première couche, on peut ajouter des services qui peuvent être payants comme les epi-journaux, les méga-journaux les services de données et d'impact. Tout ceci en permettant un accès ouvert basique gratuit pour permettre aux pays d'Amérique latine plus de participation.

Ainsi en Amérique Latine, du fait du pouvoir d'achat et de la situation économique, sont nées des initiatives Open Access centralisées puissantes, avec un fort lectorat sur lesquels on peut aussi réfléchir et noter qu'elles ont été créées avec des moyens d'Etat.

Il s'agit de plateformes qui en plus d'être des archives ouvertes proposent l'hébergement de revues.

Présentation



Amérique Latine
une tradition de libre accès non commercial

@dominiquebabini

Colloque Dynamiques de l'Édition scientifique, de l'Industrie de l'Information, de la Documentation - Un Agenda 2015 pour la science publique ouverte
Atelier 1 : Quelles exigences communes pour renforcer l'Édition scientifique ? Quels partenariats pour l'Open Access et pour l'Open Process ?
http://www.cnrs.fr/dist/z-outils/documents/MEUDON_4-5nov2014.pdf
CNRS – DIST, Meudon - 4 et 5 Novembre 2014

| www.clacso.org |



Présentation de Marc Minon

L'arbre de décision de l'Open Access

Le modèle économique de Cairn est basé sur la souscription (dépense d'abonnement).

On part d'un terrain d'observation : les publications de SHS de langue française. Le contexte est très différent de celui de SCOAP³ car en SHS Cairn ne travaille pas sur 7 revues mais plutôt sur 1600 revues.

De plus, dans le domaine des SHS, les travaux de recherche ne se diffusent pas principalement dans des revues mais surtout dans des ouvrages collectifs et des monographies. Ethiquement, politiquement, culturellement, scientifiquement, se poser la question de l'Open Access en la centrant sur les revues n'a que peu de pertinence dans ce cas.

Pour Cairn, l'Open Access se décline selon beaucoup de façons : certaines publications sont en libre accès immédiatement, d'autres en accès payant pendant une durée définie puis en libre accès après. Puis d'autres en accès payant.

Quel est le degré de proximité entre HAL et Cairn? Parmi les articles de revues diffusées à ce jour sur Cairn, moins de 5% sont déposés sur HAL. Il y a plus d'articles de revues aujourd'hui sur Academia ou sur ResearchGate que sur HAL.

Situation à ce jour en ce qui concerne les politiques d'accès aux revues : 1/3 des revues sont conformes aux recommandations de la Communauté européenne (diffusion dans les 12 mois).

Pourquoi ce contraste ?

1. Les différents acteurs échangent trop peu entre eux. La preuve dans ce colloque, on peut constater la quasi-absence d'éditeurs de SHS de langue française.
2. Il y a un manque de visibilité sur le jeu des acteurs.

Il y a différents types de solution pour l'Open Access qui répondent à un certain nombre de questions qu'on se pose séquentiellement.

Question 1 : Une diffusion plus ouverte des résultats des travaux des chercheurs est-elle souhaitable ?

Oui pour la plupart

Question 2 : Le travail éditorial (sélection, relecture, mise en forme, traduction, "fabrication", etc.) apporte-t-il une valeur ajoutée à la simple publication (sur le web) des résultats des travaux des chercheurs ?

Un certain nombre d'acteurs peuvent penser que **NON**, il faut renverser le paradigme de l'édition: sélection : ordre, amélioration, puis diffusion. Aujourd'hui le web permet la diffusion puis on peut laisser l'intelligence du réseau améliorer et hiérarchiser.

Cairn pense que le travail éditorial continue à apporter de la valeur, et doit continuer à intervenir en amont

Question 3 : Ce travail éditorial entraîne-t-il des coûts, y compris dans un environnement numérique ?

Cairn pense que **OUI**, même dans un environnement totalement électronique. Bien sûr le coût marginal pour l'utilisateur est nul sur le net mais il reste un coût moyen qui plutôt que de se réduire, se renforce d'années en années.

Question 4 : Est-il envisageable/souhaitable que ce travail éditorial soit pris exclusivement en charge par des maisons d'édition publiques, qui s'appuieraient exclusivement sur des subventions publiques ?

Un certain nombre d'acteurs pensent que **OUI**, le numérique est l'occasion pour la communauté scientifique de se réappropriier le travail capté par des maisons d'édition.

Si on pense qu'il est acceptable voire souhaitable qu'il y ait une diversité des paysages éditoriaux, la réponse est **NON**.

Question 5 : La diffusion électronique (gratuite/immédiate/non bridée) des articles et ouvrages (dans leur version finale) risque-t-elle de peser sur l'économie de ces maisons d'édition ?

Un certain nombre d'acteurs répondent **NON** : plus on diffuse sous format électronique plus on va pouvoir vendre du contenu ailleurs. Ce que l'on observe dans le domaine de la vidéo, de la presse, de la musique, va-t-on l'observer dans le domaine de l'édition ? Cairn pense que **OUI**, il y aura un impact.

Question 6 : La diffusion électronique (gratuite/immédiate/systématique) des articles et ouvrages (en version pre-print) risque-t-elle de peser sur l'économie de ces maisons d'édition ?

On peut avoir des avis très différents. Sur ce point, la réponse doit être modulée selon les disciplines concernées.

Question 7 : Les éditeurs SHS (de langue française) réalisent-ils une marge telle qu'ils pourraient supporter sans difficulté majeure une baisse de leurs revenus ?

Question 8 : Serait-il possible/souhaitable, pour les maisons d'édition, de "libérer" systématiquement le contenu des articles et ouvrages et de se rémunérer autrement ?

Etc.

Présentation



Présentation de Gemma Hersh

Sharing and access : Gold and Green open access at Elsevier

La scène de l'Open Access est très complexe et différente selon les régions et pays du globe. L'Angleterre penche pour un Open Access Gold, H2020 prône un savant mélange d'Open Access Gold et Green.

Aux Etats-Unis, les agences de financement poussent vers l'Open Access Green.

Elsevier a mené une enquête en 2013 auprès de communautés de chercheurs sur la compréhension de l'Open Access. Il en résulte que 75% ont une compréhension, voire une très bonne compréhension de l'Open Access.

De même, en ce qui concerne leur préférence de publication, il n'y a pas de nette préférence entre les différents modèles de revues (revues avec abonnement ou revues en libre accès, ou hybrides..).

Aujourd'hui la majorité des articles est publiée sous abonnement, la part d'articles en Open Access a augmenté de 20% en 2013, mais ne représente que 8% des publications.

Elsevier souhaite donc que le chercheur ait le choix et propose plusieurs politiques de publications.

La politique Open Access d'Elsevier soutient les 2 voies d'Open Access, la voie dorée et la voie verte.

La politique gold (110 revues) est sans cesse améliorée en lançant de nouvelles revues en Open Access, en mettant en place pour les revues qui ne sont pas encore en Open Access des options pour du Gold...

Le libre accès vert : Elsevier met en place des licences permettant aux auteurs de partager ou d'archiver selon plusieurs modalités. Le manuscrit accepté peut être copié/collé sur leur blog ou le site web personnel ou encore dans l'archive nationale mais parfois il ne peut être accessible en dehors de l'institution qu'après une période d'embargo.

Les nouveaux outils de partage au sein d'Elsevier

Share Link : accès pendant 50 jours à Science Direct

L'enquête Elsevier, montre aussi que les auteurs vont chercher l'information sur différents sites ou plateformes, recherche qui dure en moyenne 2h et demie. L'auteur préfère avoir la dernière version de l'article.

Les solutions Elsevier :

L'alimentation des métadonnées : Elsevier a créé **2 API l'une sur ScienDirect et l'autre sur SCOPUS** qui permettent d'intégrer des métadonnées structurées dans l'archive institutionnelle, ce qui facilite les recherches et permet la mise à disposition via ScienceDirect de l'abstract.

Les API sont étendues avec un plug-in. Dans le cas du libre accès, l'embargo est important pour maintenir le business model mais il est difficile pour les chercheurs de connaître l'embargo lié à l'article et les licences des éditeurs. Lors de l'utilisation de l'API les métadonnées concernant l'embargo sont également concernées. Et donc dès la fin de la période d'embargo, l'article sera automatiquement en libre accès.

Les fonctionnalités embarquées :

Le lecteur reste sur sa plateforme d'origine (l'archive institutionnelle par exemple) mais accède à l'article via la plateforme Elsevier.

Présentation



**Sharing and Access:
Gold and Green Open Access at Elsevier**

Gemma Hersh, Policy Director

Empowering Knowledge

Conclusions de l'atelier 1

Nouveaux modèles économiques, nouveaux process pour l'Open Access

L'Open Access (ou libre accès) n'est pas la gratuité. Le libre accès (avec ou sans embargo) est une donnée acquise, il s'agit d'établir comment y parvenir.

Il est souhaitable de garder une diversité des acteurs et des possibles, répondant à la diversité des disciplines et des méthodes de recherche. Il s'agit de se rappeler que la connaissance est un bien public.

L'Amérique latine, nous donne l'exemple. Comme dans les systèmes publics de santé et scolaire: elle considère que tout le monde a droit d'accès à cette connaissance. Ainsi, la connaissance en tant que contenu brut est en libre accès. Et il y a des niveaux supplémentaires de paiement pour les services environnants. C'est un modèle Freemium.

Pistes pour l'agenda 2015

1. Encourager l'innovation éditoriale et l'innovation économique

Il y a 3 modèles économiques du numérique :

- Freemium
- Publicité
- Souscription (unglue.it, knowledge unlatched)

Copernicus a montré des exemples d'innovation éditoriale avec l'open peer review. Il est également souhaitable de mobiliser les bibliothèques dans l'acquisition des ressources Open Access : aller vers des modèles de souscription ou de financement aval. En effet, le modèle Auteur/Payeur, n'est pas transposable à toutes les disciplines.

2. Faire passer l'expérience de nouveaux modèles économique de certaines disciplines vers d'autres disciplines (SCOAP³ a indiqué les 3 caractéristiques qui ont permis son montage du modèle économique (petite communauté à 7-8 revues, communauté green, et existence d'un organisme mondial qui porte le projet)
3. Bibliographie de l'existant : modèles et communautés en France et à l'étranger.

Atelier 2
Stéphane Pouyllau

Quels partenariats pour la construction des Plateformes Multi-Usages ?

Intervenants	Appartenance
Patrice BELLOT	Open Edition Lab
Christine BERTHAUD	CNRS-CCSD
Thierry BOUCHE	MathDoc
Jean-Baptiste de VATHAIRE	CAIRN
Olivier DUMON	ELSEVIER
Francis LEMAITRE	AAR
Stéphane POUYLLAU	CNRS-Huma-num
Yvan STROPPIA	Run&Share

Introduction

Lors de cet atelier ont été présentées des plateformes multi-usages qui, dans le cadre de l'Enseignement Supérieur et de la Recherche, permettent de gérer, de manipuler et de rassembler des métadonnées et des données de la Recherche.

Dans le monde numérique d'aujourd'hui, un nombre important de plateformes multi-usages sont créés chaque jour dans les laboratoires, les institutions ou encore chez les éditeurs. Les plateformes sont au cœur des usages futurs. Au cœur de la problématique des plateformes multi-usages un point de vigilance apparaît : la qualité des données et des métadonnées.

Le développement de plateformes, la diffusion de données, facilités par des outils, se font de manière fonctionnelle mais la qualité des données et des métadonnées est au centre de la préoccupation des chercheurs (référentiels, liste d'autorité...). Les données et métadonnées doivent être de bonne qualité pour que toutes les plateformes fonctionnent à un niveau d'exigence scientifique incontestable par les utilisateurs.

Intervenants et présentations :

- **Yvan STROPPIA** de Run&Share : Présentation de la **plateforme Execandshare** – Reproductibilité, libre accès et lien entre: articles, données de l'expérimentation et le code de mise en œuvre ;
- **Francis LEMAITRE** de la fondation des Sciences de l'Homme – Laboratoire EsCom : Présentation de la **plateforme ASA** (Atelier de Sémiotique Audiovisuelle) - Développement d'un environnement de travail qui se superpose à la bibliothèque audiovisuelle AAR et rend possible l'analyse de vidéo et du discours scientifique dans les vidéos ;

- **Thierry BOUCHE** – Cellule Mathdoc, documentation mathématique : Présentation de la **plateforme CEDRAM** (Centre de diffusion de revues académiques mathématiques) – Permettre une large diffusion des volumes courants en mathématique, aide pour la production aux meilleures normes d'édition électronique à l'archivage pérenne des articles ;
- **Patrice BELLOT** – Professeur de mathématique à Marseille : Présentation de la **plateforme OpenEdition Lab** – Introduire dans l'écosystème d'OpenEdition des innovations permettant l'exploitation des possibilités offertes par le numérique ;
- **Jean-Baptiste DE VATHAIRE** – Cairn.info : Présentation de la **plateforme Cairn.info** – Outils et services proposés sur la plateforme ;
- **Christine BERTHAUD** – CCSD : Présentation de la plateforme HAL – Dépôt en archive ouverte. Traitement, gestion pour la qualité des données et métadonnées ;
- **Olivier DUMON** – ELSEVIER : Présentation des **plateformes Scopus et Sciences Direct** : Développement et offre de nouveaux services. Travailler en collaboration avec des institutions de recherche pour le développement de nouveaux outils.

Ces présentations permettront d'obtenir une vue globale sur l'environnement des plateformes multi-usages par un mélange des disciplines et des finalités de chacune (plateformes de données de la recherche, plateformes d'édition électronique...) avec en point de focus la qualité des données et des métadonnées.

Lors de ces présentations aucune différenciation n'a été faite entre publication scientifique et données de la recherche. Stéphane POUYLLAU a souhaité ne pas faire une typologie par type de données mais par les outils et les services proposés par les plateformes, pour améliorer la qualité, la performance, pas seulement en terme de moteur de recherche mais plutôt en terme de qualité de contenu.

<p>Présentation de Yvan Stroppa La plateforme de calculs en ligne ExecAndShare</p>

<http://execandshare.org>

Cette plateforme est hébergée par la TGIR Huma-Num. Le nom Execandshare a pour origine :

- Exe = exécution en ligne
- Share = partage

L'objectif de la plateforme Execandshare est de permettre le libre accès et le lien entre les articles, les données d'une expérimentation et le code de mise en œuvre afin d'améliorer les différents échanges.

La plateforme a une architecture de calcul orientée Web, ciblée sur le concept de sites compagnons pour :

- Des articles scientifiques (working paper)
- Et des articles déjà publiés.

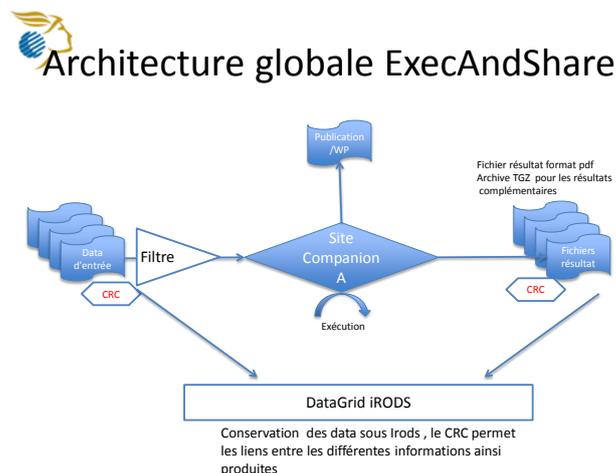
Cette plateforme est née dans un contexte économique avec au départ des partenariats avec des grandes écoles comme HEC. Execandshare est née de plusieurs constats : de plus en plus de codes informatiques viennent enrichir les publications scientifiques mais peu de ces codes sont diffusés ou accessibles (2/3 ans) et, lorsqu'ils le sont, ils sont souvent trop complexes. L'idée est donc de faciliter la lecture du code. Cette plateforme a vocation de mettre en ligne les codes et d'en faciliter l'exploitation, la lecture.

Les objectifs

- **Diffuser** de façon rapide et à l'échelle mondiale les résultats de la recherche académique
- Permettre à une large communauté d'utilisateurs (chercheurs, étudiants, professionnels...) d'**utiliser** les méthodes scientifiques les plus récentes
- Fournir à la communauté académique (chercheurs, éditeurs, arbitres) un moyen de pouvoir **répliquer** les résultats scientifiques ; Grâce à une utilisation simplifiée des codes les chercheurs pourront répliquer un résultat de calcul qui sera ensuite utilisé pour créer des algorithmes spécifiques qui seront à leur tour analysés.

Ces procédés permettront également aux chercheurs de vérifier le résultat d'une recherche à partir du code présent dans la publication.

Architecture globale



Les données d'entrée qui passent par des filtres sont analysées pour être traitées par le site compagnon. Le site compagnon permet de faire le lien avec la publication dans laquelle se trouve le code.

Toutes les données échangées sur Execandshare sont conservées grâce à l'utilisation du DataGrid IRODS.

Inputs

Les données sont appuyées sur un CRC calculé à chaque entrée. Toutes ces entrées sont définies et décrites par le chercheur, l'utilisateur. Aujourd'hui, l'ensemble de ces données ne sont pas encore associées à toutes les informations qui pourraient permettre de caractériser la donnée. Ce volet est en cours de développement.

Site compagnon

Le site compagnon est composé de fichiers sources à partir desquels sont construits les exécutables (travail préparatoire) puis un travail est mené au niveau du code lui-même. Les liens du site compagnon se font selon les mots clés associés. Par ailleurs, des descriptions algorithmiques permettent de décrire la composition du code.

Outputs

Les résultats de calcul sont présentés sous forme de rapports PDF. Si le code est amené à générer plus de calculs, ceux-ci sont conservés dans un fichier archive et seront taggués au niveau d'IRODS. La notion de traçabilité est importante pour conserver et définir l'information qui a permis d'obtenir le résultat. Les formats sont condensés pour permettre l'accès à tous les calculs, les procédés qui ont permis de former le résultat. Execandshare est également une plateforme collaborative qui permet à des utilisateurs d'enrichir les données, les informations, avec une possibilité de versioning. Les codes peuvent donc évoluer dans le temps. Pour permettre de faire vivre le code il est nécessaire de garder le lien initial mais aussi de permettre le versioning. Le lien de l'ancienne version est conservé et le lien vers le nouveau code est ajouté. Ce procédé permet de conserver tout le contenu produit sur un site compagnon et ainsi permet de créer un axe temporel qui permet des comparaisons et l'analyse des évolutions du code.

Création d'un site compagnon

La création des sites compagnons est faite par l'auteur, le chercheur avec ses co-auteurs (circulation entre les différents auteurs qui ont permis de créer le code) puis la mécanique une fois validée par tous les auteurs, une intégration manuelle est effectuée par l'équipe d'Execandshare qui met le site en production. Le procédé se déroule ainsi :

- **Les phases de l'auteur** : l'auteur ne diffuse pas l'article dans lequel se trouve le code mais uniquement le lien vers le site de l'éditeur ou autre entrepôt de données dans lequel se trouve l'article, le résultat de sa recherche. L'auteur fait une description fine des inputs, et décrit le code, tout ce qui a contribué au résultat avec une description des sorties, une prévisualisation et la rédaction d'un résumé. Une fois ces étapes accomplies, l'auteur doit demander la validation par tous les co-auteurs s'il y en a.
- **Phase d'intégration** : L'équipe ExecandShare, une fois le site validé par tous les auteurs, vérifie l'exhaustivité du résultat dans les mêmes conditions que le chercheur (OS, outils, logiciels,...). Si l'équipe constate une erreur ou un écart le chercheur est interrogé. Ce procédé permet de vérifier qu'il n'y ait pas de problématiques liées au portage. Une fois le résultat validé, le site est mis en production. Un rapport d'intégration est ensuite retourné au chercheur sur les traitements effectués sur son code.

- **Site** : Les utilisateurs d'Execandshare peuvent accéder à l'ensemble des sites compagnons. Le chercheur a accès à la gestion de ses sites compagnons avec un système de FAQ qui lui permet de gérer les problématiques rencontrées par les utilisateurs du site.

Présentation



Dynamiques
de l'Édition scientifique, de l'Industrie de
l'Information, de la Documentation

Plateforme ExecandShare

<http://execandshare.org>

Atelier #2

B. Charroux - L. Sliman – Y. Stroppa



Présentation de Francis Lemaître

L'environnement ASA (Atelier de Sémiotique Audiovisuelle) du programme
« Archives Audiovisuelles de la recherche »

<http://asahs.hypotheses.org/>

L'environnement ASA a été créé autour des archives, des contenus du programme d'Archives Audiovisuelle de la Recherche (AAR) lancé en 2001. Ce programme a permis la création d'une bibliothèque audiovisuelle en ligne. Depuis 2009, le programme ASA-SHS permet le développement d'un environnement de travail qui se superpose et rend possible l'analyse de vidéo et du discours scientifique des contenus AAR. Ces analyses peuvent varier selon le domaine scientifique dans lesquels elles se positionnent. Un programme ANR intitulé « **CAMPUS-AAR** » (démonstrateur d'un environnement numérique pour la production, description/indexation et publication d'archives audiovisuelles) a été lancé pour supporter ASA.

L'environnement ASA est multi-projets et multipartenaires. L'objectif est de permettre l'exploitation des fonds médias qui proviennent de plusieurs contenus : images, textes (sauf la

partie segmentation qui n'est accessible que pour les vidéos), etc. L'environnement est construit selon un domaine particulier mais aussi des objectifs particuliers.

Un outil « Sémioгене », commun à l'ensemble de l'environnement, permet la gestion des médias. Une fois le média chargé par Sémioгене il peut être exploité. Pour décrire le média (segmentation et description sémiotique), l'outil « Interview » est utilisé pour la segmentation des vidéos. L'outil « GAAV » (atelier de publication AAR) permet de rajouter des informations ou éditer du code HTML (cela dépend de l'ancienneté de l'environnement utilisé). Grâce à ASA on peut à partir d' « Interview » re-sélectionner un média, et faire des analyses spécifiques. L'outil « Sémioscape » permet la publication ou la republication de dossiers thématiques.

Plusieurs modèles de données se trouvent sur ASA :

- **Médias** (vidéo, image, PDF) : stockés sous forme de notices ;
- **Analyses** : description appropriée d'un média et des parties d'un média en fonction d'un domaine de connaissances et des objectifs visés par une description. L'interfaçage permet de guider l'analyse ;
- **Ressources métalinguistiques** : structure du sujet, modèle de description, basées sur des thésaurus.

Média

Pour les données média sous forme de fichiers, chacun d'eux contient des informations techniques. Les données média sous forme de notice utilisent un vocabulaire contrôlé dans le thesaurus développé au format propriétaire pour gérer des thésaurus externes.

Actuellement les médias sont stockés en SQL et extraits en XML. Un projet est en cours pour donner accès aux formats des plateformes utilisés par HAL, Isidore, le dépôt légal de la BNF.

Analyse

L'analyse se rapporte à un média et il n'y a pas de report des informations déjà liées au média. Les trois types d'analyses se font :

- Au niveau de la méta-description (niveau le plus haut)
- Au niveau du document
- Au niveau du segment (niveau le plus bas).

Analyse des acteurs

Une analyse du rôle des acteurs dans les vidéos peut être effectuée (acteur, auteur, caméraman...). Il est également possible de décrire les plans visuels et sonores et les techniques visuelles associées.

L'analyse du contenu est pilotée en fonction du domaine de recherche. Les descripteurs dans le thésaurus sont alignés avec RAMEAU.

Le projet Campus AAR

Ce projet ANR (« *Digital Humanities* ») a débuté le 1^{er} Janvier 2014 et prendra fin le 31 Décembre 2016.

Les partenaires associés sont :

- Fournisseurs de contenu :
 - **FMSH - ESCoM-AAR** (coordinateur) : Archives Audiovisuelles de la Recherche
 - **CERIMES** : Canal-U
 - **CNRS-CCSD** : HAL Vidéo
- Partenaires techniques :
 - **INA** : modèles de données, outils d'analyse
 - **ARMADILLO** (industriel) : bases de données, outils de publication

Site web: <http://campusaar.hypotheses.org>

Présentation



ENVIRONNEMENT ASA
(ATELIER DE SÉMIOTIQUE AUDIOVISUELLE)
UNE PLATEFORME SCIENTIFIQUE ET TECHNIQUE DES ARCHIVES AUDIOVISUELLES DE LA RECHERCHE

Francis Lemaitre
Fondation Maison des Sciences de l'Homme (FMSH, Paris, France)
Equipe Sémiotique Cognitive et Nouveaux Médias (ESCOM-AAR)
Archives Audiovisuelles de la Recherche (AAR)
lemaitre@msh-paris.fr

<p style="text-align: center;">Présentation de Thierry Bouche Le circuit des métadonnées à travers les plateformes de mathématiques</p>

<http://www.cedram.org/>

Le CEDRAM est un projet soutenu par l'INSMI (CNRS) et l'Université Joseph FOURIER de Grenoble. Le projet soutient les revues académiques de recherche en mathématiques. La réalisation se fait à la cellule Mathdoc de Grenoble.

Ce projet regroupe :

- Un ensemble d'outils de production
- Un portail de diffusion www.cedram.org

L'élargissement du périmètre et du nombre de services est en projet.

Les documents contenus dans le CEDRAM ne sont pas soumis à des modifications avec versioning. Une fois publiés les textes ne sont jamais modifiés. Une revue ne sera pas retenue par le CEDRAM si l'auteur/l'éditeur n'adhère pas à la politique de barrière mobile. Les revues rendues disponibles dans le CEDRAM sont en accès libre après une période de temps défini.

Deux fonctions à ce projet :

- Fonction publication : produire des nouveautés
- Fonction bibliothèque : conserver les nouveautés

La publication est un instant dans la vie d'un article et l'objectif du CEDRAM est de faire partie d'un réseau qui suit la vie de l'article. Les sources des articles sont pour le moment archivées dans un système de sauvegarde pour leur préservation. Les données du CEDRAM sont dupliquées dans la base de données NUMDAM après 1 à 2 ans (métadonnées et textes). NUMDAM est partenaire de EuDML et bientôt GDML pour une plus grande visibilité et la construction d'une bibliothèque de références couvrant toute la discipline.

LE CEDRAM a constitué son propre programme de production, une chaîne de production maîtrisable par de petites structures avec uniquement des formats en entrée de type Latex, bibtex, etc. Cette chaîne de production permet de produire des métadonnées exactes. Les formats pivots ne sont pas des fichiers de type XML mais plutôt PDF (déterminés par les sources Latex). Avec les sources Latex il est possible de connaître exactement les données utilisées par le chercheur. Tout le contenu du CEDRAM est généré à partir de sources TEX.

Le CEDRAM est encore loin d'une bibliothèque universelle en mathématique (seulement 6 % des ressources acquises). Grâce à son partenariat avec EuDML, qui participe à l'assemblage de plusieurs projets de numérisation, (assemblage de 200 milles articles et tests de technologies) et des financements européens, le projet espère se développer au niveau européen et propager ses idées à l'international.



Le circuit des métadonnées à travers les plateformes de mathématiques

Thierry Bouche



*Dynamiques de l'édition scientifique,
de l'industrie de l'information,
de la documentation*
Meudon, mardi 4 novembre 2014

cedram

EuDML

NUMDAM

Présentation de Patrice Bellot

Open Edition et la fouille des textes : de l'analyse automatique de références
bibliographiques à l'aide à la recommandation de lectures

L'environnement Open Edition Lab a été créé après obtention d'un Equipex qui a permis d'augmenter le nombre de livres, revues, disponibles sur la plateforme Open Edition. L'objectif du lab est de faire en sorte que le traitement full text, le TDM, soit possible sur des ressources de mêmes formats.

L'angle d'attaque est le traitement de données massives : qu'est-ce qui est automatisable ? Jusqu'où ? Par qui ? La réflexion est portée sur des termes d'application avec pour finalités le développement de fonctionnalités de lecture et d'écriture ainsi que de systèmes de navigation et de recommandation.

Pour cela deux axes sont exploités :

- La fouille de texte
- L'innovation éditoriale

Fouille de texte (Text & Data Mining – TDM)

Les objectifs sont le développement de services à forte valeur ajoutée ainsi que de systèmes de recommandation s'appuyant sur les usages, les contenus et les liens entre les documents.

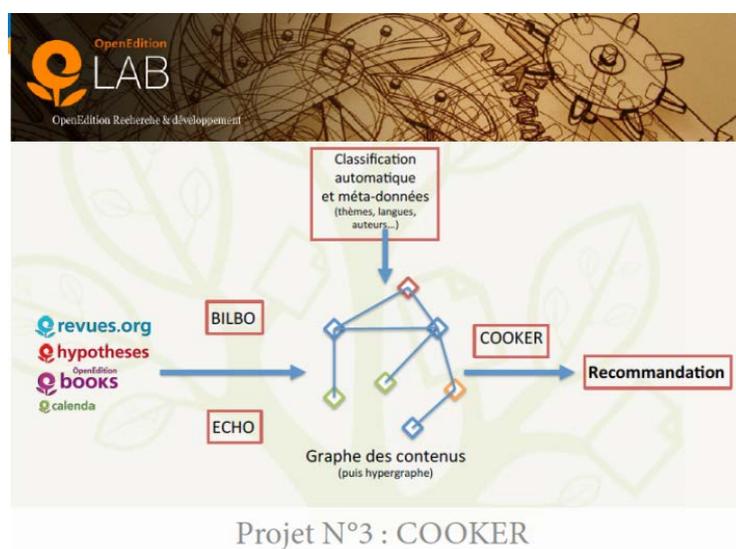
L'objectif est de relier les contenus entre eux afin d'aider la navigation du chercheur. Pour cela, il faut travailler sur les référencements qui permettent d'ajouter des métadonnées orientées. Exploiter ces enrichissements permet d'exploiter les méta-informations autour des données.

Le premier projet de recherche et développement d'Open Edition lab, BILBO, a été lancé en 2011. Ce logiciel est basé sur des approches d'apprentissage automatique et permet de détecter et annoter sémantiquement les références bibliographiques. Pour cela, BILBO identifie les éléments qui composent la référence, enrichit l'encodage et relie la référence à son DOI.

BILBO est disponible pour des tests et le logiciel est en Open Source.

D'autres projets ont par la suite été lancés et viennent en complément de BILBO :

- Le projet ECHO : pour la détection automatique de compte-rendu de lecture
- Le projet COOKER – Projet de collaboration basé sur la recommandation au travers d'un ensemble de critères basés sur une liste de plusieurs milliers d'entrées. Cette liste est utilisée pour l'ensemble du référencement



Présentation



Présentation de Jean-Baptiste de Vathaire Métadonnées et intertextualité : favoriser la découverte par l'interopérabilité et le balisage sémantique

La plateforme Cairn effectue un travail d'agrégation sur les contenus d'une centaine d'éditeurs et structures éditoriales. La structuration des métadonnées permet de favoriser la découverte et les rebonds entre différents articles dans un corpus important de plus de 300 000 articles qui touchent toutes les disciplines SHS. Dans cette optique le choix de Cairn a été de structurer tous ses contenus.

Cairn a signé la charte de revues mobiles qui lui permet de rendre disponible en libre accès plus de 60% des contenus sur sa plateforme.

L'objectif est de faciliter l'accès par la structuration des données, l'indexation, le service OAI, des outils de découverte, etc.

Interopérabilité

La plateforme Cairn a signé des accords d'interopérabilité avec d'autres portails, sites, etc. La structure, la qualité des métadonnées sont essentielles dans cette démarche car elles permettent la continuité d'accès.

De même, avec la syndication de contenu (une partie d'un site est accessible depuis d'autres sites), la plateforme doit être capable de récupérer rapidement l'ensemble des contenus accessibles. Pour cela, il apparaît nécessaire de structurer et de permettre les traitements sémantiques sur les contenus.

Mise en place d'un outil « autocompleted »

La mise en place de cet outil a permis à la plateforme Cairn un élargissement sémantique de la recherche. Ce système se base sur un balisage sémantique, un traitement du texte. Même lorsque la recherche est tapée en anglais, l'élargissement sémantique, par une réorientation/traduction, permet de faire ressortir les contenus francophones. L'objectif de cette outil est d'intéresser les chercheurs étrangers à la recherche francophone.

Analyse matricielle

Un outil d'analyse sémantique des termes pertinents est nécessaire pour la création d'un calcul matriciel du sens global des contenus et l'affichage des contributions (citation, extraits de revues ...) qui sont sur un seuil de proximité permettant les liens de co-citations.

Dans le domaine des SHS, le « cité par ... », le référencement, peut s'avérer difficile en raison d'un manque de règles de référencement, de différences typographiques.

Enrichissement sémantique

La nécessité d'étendre et d'industrialiser l'enrichissement sémantique a amené Cairn à acquérir la solution Luxid de Temis ce qui lui a permis de bénéficier de modules déjà développés et de cartouches de connaissance déjà construites. Temis a donné à Cairn la possibilité de tester ses outils afin de vérifier si ceux-ci étaient adaptables à la plateforme Cairn.info. La solution Luxid a permis à Cairn d'obtenir, sans perte de qualité sur la reconnaissance, 2 millions de références reconnues et identifiées ainsi que 300 000 liens DOI. A partir de Luxid, les perspectives de développement apparaissent nombreuses car la base permet de travailler sur d'autres éléments pour la création de valeur à partir d'éléments peu ou pas structurés.

Présentation



Cairn.info

Métadonnées et intertextualité

*Favoriser la découverte par la structuration
et le balisage sémantique*

Meudon
Mardi 4 novembre 2014

Cairn.info
Chercher : Repérer : Avancer

{ 04/11/2014 }

Présentation de Christine Berthaud Processus qualités et référentiels dans HAL

HAL est une plateforme d'archive ouverte qui permet l'archivage et le dépôt de pré-prints ou d'articles publiés.

Réseau

Pour le CCSD, l'un des premiers composants de la qualité de la plateforme HAL est le réseau. Tous les établissements qui ont signé la convention HAL ont obligation de déposer dans HAL. HAL a été développé en 2001 et d'autres plateformes qui y sont connectées ont été créées comme Sciencesconf.org (organisation de congrès) qui impose le dépôt des actes des évènements organisés dans HAL.

Les caractéristiques de HAL

- Bilingue
- Choix de plusieurs disciplines
- Connexion avec des archives importantes : ArXiv, Open Aire ...
- Archivage pérenne : duplication de la base d'HAL et procédure d'archivage pérenne des documents pour rendre toutes les versions des docs visibles pour l'éternité numérique

- Une équipe dédiée : de développeur, de professionnel de l'IST, de management de communauté
- Hébergement au centre de calcul de l'IN2P3 : garantie de sécurité, réseau performant, accès H24

La V3.0 de la plateforme a été mise en ligne très récemment, la qualité passe donc aussi par le « débogage » de cette nouvelle version. Dans la V3, il est à présent possible de déposer des médias, des données.

Authentification Centralisée

HAL permet aux institutions de créer leur propre archive. Bientôt tous les utilisateurs de HAL pourront accéder à tout le contenu à travers un compte unique.

Un nouvel identifiant HAL pour les chercheurs a été créée. Celui-ci est relié facilement aux identifiants les plus communs comme ORCID ou ResearcherID. Dans le futur, des connexions seront également possibles avec le système d'identification de RENATER.

Les référentiels

Les référentiels sont les garants de la qualité. Tous les laboratoires, tous les établissements, toutes les institutions de l'ESR, sont présents dans le référentiel de HAL. Autour du référentiel existe un travail de mise à jour quotidien à la demande des laboratoires, des auteurs. La qualité dépend de ce travail. Le référentiel auteur de HAL est composé de plus de 700 000 auteurs (pas uniquement français). Pour la nomenclature des domaines scientifiques, HAL avait fait le choix de s'aligner sur ArXiv mais avec la création de son propre référentiel basé sur ArXiv. Les référentiels des contrats ANR, des saisis et d'une liste de revues ont été intégrés tout comme des grands thésaurus internationaux.

Contrôle de la qualité des dépôts

La qualité passe aussi par un control de qualité sur les dépôts. Lorsque l'auteur dépose il y a une vérification systématique humaine sur le document déposé. L'importance est de pouvoir toujours citer un document, il y a donc en ligne plusieurs versions d'un même article (le pré-print, l'article, l'article modifié, les données, les médias) avec un identifiant pérenne et citable. Pour préserver le document celui-ci est lié à une URL stable qui ne dépend pas du stockage physique du document.

En conclusion, le travail effectué sur HAL permet de maintenir le lien avec les grandes archives internationales et disciplinaires. Il est essentiel de permettre la cartographie de la carrière d'un chercheur, de la production d'un chercheur, de rendre publique une liste de la production d'un chercheur sur une plateforme ouverte. Enfin, HAL travaille également sur les statistiques concernant les usages en matière de dépôt en archive ouverte ; à partir d'un document de nombreuses données peuvent être croisées. Il est possible d'accéder à HAL par Google mais ensuite un moteur de recherche très performant, du même type que celui utilisé par Isidore, permet une recherche très riche sur les archives ouvertes.



Référentiels et qualité

christine.berthaud@ccsd.cnrs.fr



Présentation de Olivier Dumon

Métadonnées et données de recherche : structuration, management, interopérabilité, ouverture et transparence

Données et métadonnées

Chez ELSEVIER sont développées 30 ontologies, pipelines d'annotation, sur la plus part des plateformes utilisées et dans 10 des solutions proposée par l'éditeur.

Text & Data Mining (TDM)

A travers son API, Elsevier rend accessible tous les articles de Science Direct (\simeq 12.5 millions) en full Text. Le TDM est libre sous licence CC-BY-NC pour les abonnés. Pour les chercheurs moins à l'aise avec le TDM, Elsevier a développé en collaboration avec *The National Centre for Text Mining* (NaCTeM) un procédé s'appuyant sur les technologies du Cloud évitant la nécessité pour les chercheurs de construire et entretenir les infrastructures TDM.

Les données de recherche

Elsevier souhaite la mise en place d'un maximum de liens entre les publications des articles. Pour cela la plateforme est dotée de widgets permettant aux auteurs de remonter leurs données sous le format de leur choix. Les citations géographiques peuvent être liées à google maps. Les

Widgets sont servis par les *data repository*. Un pilote a été mis en place pour montrer les données de la recherche en CCP.

Identifications des chercheurs

Les analyses ont montré que 50% du temps des chercheurs est consacré à des tâches administratives. Les chercheurs disposent par ailleurs de 5 et 10 profils en ligne (institutionnels, liés à des projets spécifiques, profils privés sur internet, profils sur des réseaux scientifiques, profils utilisateurs agrégés par les agences d'évaluation).

En matière d'identifiant, la base ORCID rencontre un succès important avec l'identification de plus de 800 000 chercheurs par une ORCID ID. Par l'intermédiaire de Scopus, Elsevier est l'un des principaux fournisseurs d'ORCID. Les chercheurs sont perdus dans tous ces profils ; les chercheurs veulent plus de maîtrise sur les données utilisées pour l'évaluation, pour le Benchmarking. Aucun des profils cités précédemment n'est utilisé au quotidien par les chercheurs dans leur travail.

Selon Elsevier, les chercheurs créeront des profils privés sur lesquels ils auront le contrôle et pourront déterminer les informations qu'ils souhaitent rendre public. Un prototype lié au profil privé des utilisateurs de leurs plateformes a été mis en place par Elsevier pour comprendre l'utilisation des publications des chercheurs «The author dashboard».

Elsevier souhaiterait mettre en place des partenariats avec le CNRS et d'autres institutions sur les différents sujets exposés.

Questions/discussion :

Deux éléments ressortent des présentations :

- **Construction, mise en production, exploitation de plateformes** : notion de référentiel ouvert et réutilisable, rejouable ; Circuit de mise à jour des référentiels ;
- **Qualité des données et métadonnées** : une barrière doit être franchie en matière de qualité pour la réutilisation, le traitement, l'enrichissement des données et des métadonnées.

Discussion autour de l'identification des chercheurs : ORCID

Au Portugal, une volonté politique a souhaité la création d'un profil ORCID pour tous les chercheurs. Faut-il mettre en place une obligation de création d'identifiant sur ORCID pour les chercheurs en France ? Faut-il inciter les chercheurs français à créer un identifiant ORCID par la contrainte ou par un service ?

L'incitation en France pourrait ne pas être la bonne approche. La solution pourrait être, comme dans HAL, de lier les comptes créés sur les plateformes aux identifiants communs (ORCID, ResearcherID...). Par ailleurs, il faut éduquer, former les chercheurs pour la gestion de leur identité numérique. Le CCSD mène des actions pour la diffusion d'information auprès des chercheurs sur l'importance de l'identité numérique.

Conclusions

L'ensemble des plateformes présentées lors de l'atelier 2 rencontre des problématiques de qualité pour le fonctionnement des traitements, des enrichissements de leurs contenus, pour la cartographie, pour générer du contenu réutilisable.

En termes d'ouverture des référentiels, le sentiment est que les chercheurs, les utilisateurs des plateformes présentées souhaitent échanger, partager sur ce sujet. Une discussion pourrait être mise en place autour de GeoNames en SHS pour son utilisation dans le domaine scientifique. La base de données géographique GeoNames couvre tous les pays et continents avec plus de huit millions de toponymes qui sont disponibles gratuitement au téléchargement.

Par ailleurs, tous les intervenants de cet atelier ont manipulé du vocabulaire lié à l'information scientifique et technique (IST), la création d'un lexique du vocabulaire de l'IST, un référentiel des référentiels de l'IST pourrait s'avérer intéressante.

Enfin, les professionnels de l'IST et les chercheurs doivent travailler ensemble à la mise en place de fiches pratiques pour informer les communautés scientifiques sur les enjeux des référentiels et de l'identité numérique.

Restitution de l'atelier 2

Le plan de l'atelier 2 avait pour volonté d'insuffler la question de la qualité des données et métadonnées dans la construction des plateformes multi-usages.

Les présentations ont permises d'éclairer certains points qui donnent corps aux problématiques des attendus en matière de qualité des données et des métadonnées. La question sous-jacente était : quelles exigences doit-on avoir en termes de qualité des données lorsqu'on construit des plateformes multi-usages ?

Les présentations

Lors de ces présentations il a été constaté que le terme « multi-usages » traite d'un concept qui mixe les données de la recherche et les résultats (publications). Les résultats pouvant être des objets multimédias ajoutés aux contenus des publications de manière perpétuelle et qui doivent être conservés, dont on doit conserver l'usage. Les présentations ont également beaucoup mentionné la production d'enrichissements (sémantiques, informationnels, sociaux). Cette production pose un certain nombre de questions sur les usages quotidiens des chercheurs.

Enjeux / proposition pour l'agenda 2015 :

- **Référentiels communs** : Auteurs, lieux, matières, disciplines, etc.

Proposition : Mise en place d'une politique nationale d'exposition, de valorisation des référentiels.

- **Lexique** : De quoi parle-t-on ? Stabilisation de la façon de nommer les choses afin de permettre de meilleurs échanges.

Proposition : Création d'un lexique, d'un dictionnaire de l'IST.

- **Enrichissements** : quelles valeurs ? Pour qui ? Comment est mesurée la qualité des enrichissements réalisés ? Quelle valeur pour le travail scientifique ? Analyser la valeur des enrichissements dans les travaux des chercheurs.

Atelier 3
Bruno David et Francis André

Quelles dynamiques pour valoriser l'information et la publication scientifique ?

Intervenants	Appartenance
Francis ANDRE	CNRS -DIST
Lisa COLLEDGE	ELSEVIER
Annie CORET	CNRS-DIST
Bruno DAVID	CNRS - Inee
Olivier DUMON	ELSEVIER
Odile HOLOGNE	Inra
Marie-Christine JACQUEMOT	CNRS-Inist
Hylke KOERS	ELSEVIER
Nicolas NOQUET	THOMSON- REUTERS
Guillaume RIVALLE	THOMSON- REUTERS
Alain ZASADZINSKI	CNRS-Inist

Ordre du jour :

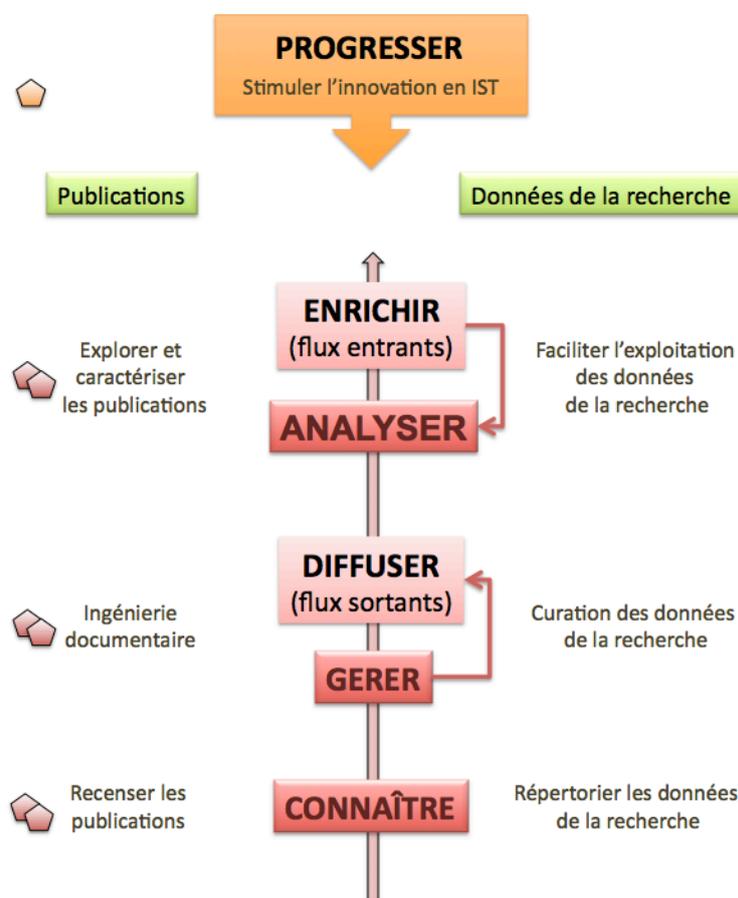
- A. Quels principes de construction des systèmes d'information sur les publications ?
- B. Quelles chaînes de valeurs afficher dans la valorisation des informations sur la publication ?
- C. Quels principes de travail afficher dans les processus d'exploration des données et publications ? Des principes de travail pour le TDM, pour la valorisation des Contenus Générés par les Utilisateurs (CGU), pour les autres formes de valorisation.

Introduction (Bruno David)

Les thèmes abordés dans cet atelier rejoignent ceux du PAP3 « Analyser et valoriser l'information scientifique » dont l'objectif central est d'accompagner les besoins d'analyse des publications et des données par des outils et pratiques innovants et partagés et d'aboutir à des services à la communauté scientifique à l'attention des chercheurs, unités, instituts et du CNRS dans son ensemble.

L'architecture globale du PAP3 identifie trois niveaux (en rouge sur la figure) qui sont symétriquement déclinés en actions pour les publications (colonne de gauche) et les données de la recherche (colonne de droite). Le premier niveau "CONNAITRE" se décline en deux actions qui sont "Recenser les publications" et "Répertorier les données de la recherche". Il s'agit de savoir ce que produisent les laboratoires et les chercheurs. Le deuxième niveau "GERER" est

étroitement associé à la nécessité de diffuser; les actions qui l'accompagnent sont: "Ingénierie documentaire" et "Curation des données de la recherche". Il faut ici être en capacité d'organiser et de gérer ce qui est produit et de le mettre à disposition (organisation des flux sortants). Le troisième niveau "ANALYSER" est l'aboutissement logique des deux premiers, il doit s'enrichir de flux d'informations entrantes qui viennent compléter, par exemple à des fins de comparaison, nos propres données. Il est assorti de deux actions: "Explorer et caractériser les publications" et "Faciliter l'exploitation des données de la recherche". Enfin, dans une perspective d'évolution et d'accroissement des services rendus, l'ensemble est chapeauté par l'obligation d'être en capacité de progresser, traduite par une septième action: "Stimuler l'innovation en IST".



L'atelier s'est déroulé en 3 temps :

- Interventions des offreurs de service
- Interventions des académiques (initiatives des institutions de recherche pour valoriser leurs productions scientifiques – publications et données)
- Discussion et pistes de travail en commun

Présentation de Guillaume Rivalle
La chaîne de traitement de la publication d'un article à sa valorisation

Les données du Web Of Science

- Le processus de sélection dans la Core collection est objectif et unique : aucune intervention d'éditeurs de revues scientifiques, 16 experts à temps plein, un taux d'acceptation de 10-12% et un suivi continu du contenu. 12700 revues sont traitées dont 10% en open access. La Core Collection du WOS est à 59% constitué de revues européennes.
- Les métadonnées sont capturées à partir des revues, les références sont traitées depuis 1900 : 58 millions de notices et 900 millions de citations. Tous les auteurs et les adresses et depuis 2008 les sources de financement sont présentes.

Enrichissement collaboratif d'une publication au niveau institutionnel et par degré de structure et référentiel interne (outil SAMPRA)

- Des règles sont définies pour alimentation du champ « Organization enhanced » à partir des adresses rendant ainsi les institutions plus visibles : 1746 pour le CNRS, 633 pour l'Inra, 68 pour l'institut Curie et ces règles peuvent être communiquées aux institutions en vue de leur amélioration (mais il faut être abonné au WoS et au JCR). 4500 institutions ont ainsi été unifiées à ce jour.
- SAMPRA a été développé par le CHRU de Lille pour permettre à un chercheur de recenser ses publications. Dans le cadre d'un partenariat CHRU de Lille et Thomson Reuters, les données du WoS sont mises à disposition gratuitement et la liste des publications validées fournie à Thomson Reuters permet d'optimiser le champ « Organization enhanced ». Les informations sur les départements ou unités d'appartenance sont également utilisées.

Valorisation et analyse bibliométrique d'un corpus de publications selon différents angles et indicateurs grâce à l'enrichissement collaboratif des métadonnées.

- La combinaison de SAMPRA avec Incites permet des analyses au niveau chercheur, unité, département.
- Incites permet de produire des indicateurs (près de 200 différents) et des rapports à tous les niveaux en fonction des données transmises à Thomson Reuters.

Commentaire :

Incites n'est pas gratuit.

Data Citation Index : faciliter la découverte, l'utilisation et l'attribution des jeux de données de recherche

- Le Web of Science est un outil de découverte utilisant les liens entre publications. Il s'agit de faire de même avec les jeux de données qui sont également des productions scientifiques, par navigation citationnelle. Il existe une multitude de dépôts de données bien distincts et isolés, si bien que leurs formats et les fonctionnalités qu'ils offrent diffèrent significativement.
- Comme pour les publications, les dépôts de données sont sélectionnés selon un processus se voulant strict et neutre.
- La mise en place de Data citation index donne lieu à la proposition d'un standard de citation des données.
- Sur 900 étudiés, plus de 200 dépôts sont indexés (critères souvent liés aux problèmes de format).

Commentaire :

Le DCI est un facteur de reconnaissance intellectuelle des données et permettra des mesures d'impact dans le futur.

Thomson Reuters préconise l'utilisation d'Orcid pour identifier les auteurs car les éditeurs les intègrent de plus en plus mais il a été précisé que l'ISNI existait au niveau international.

Présentations



Guillaume Rivalle
Responsable avant-ventes Europe
4 Novembre 2014



La chaîne de traitement de la publication d'un article à sa valorisation

Guillaume Rivalle
Responsable avant-ventes Europe
4 Novembre 2014



Présentation de Lisa Colledge
Snowball/études analytiques/services nouveaux (valorisation des institutions)

Snowball metrics a pour objectif est de mettre en place des métriques afin de permettre des prises de décisions éclairées et transparentes. Il est le fruit d'une collaboration académique-industrie.

Environ 35 métriques sont produites.

Présentation de Hylke Koers Data linking and linked data

De nombreuses données sont manipulées tout au long du processus de recherche. L'accès aux données est essentiel mais difficile.

Elsevier aide les chercheurs à stocker, partager, publier et découvrir les données.

La visualisation des données depuis l'article est possible par la mise en place de liens entre articles et données dans des repositories.

Le référentiel de données liées (Linked Data Repository) est un élément clé de la Smart solution d'Elsevier. Il fournit un ensemble de services et d'API pour les applications construites par Elsevier et des tiers pour stocker et récupérer des contenus. Le LDR stocke les données sous forme de documents RDF.

Par contre, il est nécessaire d'avoir des identifiants persistants.

Présentation

Data-linking & Linked Data

Connecting the dots in a distributed research environment



Dr. Hylke Koers, Head of Content Innovation

CNRS Meudon, 4 November 2014

Interventions « des académiques » (initiatives des institutions de recherche pour valoriser leurs productions scientifiques – publications et données)

Présentation d'Odile Hologne
Lancement d'un "data journal" en agriculture

L'incitation au partage des données est fort dans H2020 et à l'Inra mais les chercheurs ont des réticences par crainte de remise en cause de leur travail parfois, mais aussi parce que la curation et la publication des données ne sont pas reconnues dans l'évaluation.

L'une des solutions est de faire entrer les données dans la sphère des publications par l'évolution des journaux classiques (proposition de jeux de données au peer reviewing) et par les « data journal » (publication d'articles décrivant le processus d'obtention, la structure des données...).

L'Inra a créé un « data journal » mais n'est-ce pas une solution transitoire pour faciliter l'accès aux données dans un système de communication scientifique en transition ?

Présentation



Présentation de Marie-Christine Jacquemot

Le projet SIM4RDM dans lequel l'Inist a un rôle de facilitateur est un projet européen ERA-Net du 7ème programme cadre d'une durée de deux ans qui vise à permettre aux chercheurs d'utiliser de façon efficace les infrastructures techniques émergeant dans l'espace européen de la recherche en s'assurant qu'ils ont les connaissances, les compétences et les infrastructures de soutien nécessaires pour adopter de bonnes pratiques dans la gestion de leurs données.

Il a débuté par une phase d'évaluation des chercheurs dont les conclusions sont qu'il y a peu de différences entre les disciplines et que le niveau de maturité en matière de gestion des données est basique. Une sensibilisation a ensuite été mise en place. Il est aujourd'hui dans une phase pilote dont l'objectif est d'appliquer de bonnes pratiques à un projet. C'est dans ce cadre qu'une collaboration a été mise en place entre l'Inist et un OSU. Il implique un «ingénieur documentaliste» et un informaticien de l'Inist.

La méthode consiste à :

- partir d'un besoin immédiat d'intégration et exploitation de données d'observation et expérimentales
- susciter/construire une réflexion
- s'appuyer sur l'existant et travailler avec les chercheurs pour faire évoluer leurs pratiques

Présentation



The slide features a light blue background with a large, faint circular graphic. On the left, there is a dark blue circle containing the CNRS logo (the letters 'cnrs' in white) and the tagline 'dépasser les frontières' written vertically. Below the logo is the website address 'www.cnrs.fr'. To the right of the logo, the main title 'Vers une gestion et une réutilisabilité des données de recherche de l'OSU OTELO' is displayed in white text. Below the title, the presenter's name 'Jacquemot-Perbal Marie-Christine, Inist-CNRS' is listed. At the bottom, the event details 'Colloque Edition Scientifique Meudon, 4 et 5 novembre 2014' are provided.

<p style="text-align: center;">Présentation de Annie Coret et Alain Zasadzinski Conditor, vers un référentiel national des publications scientifiques</p>

Conditor répond à une recommandation de BSN3 (signalement) : construire un référentiel de la production scientifique française en s'appuyant sur des référentiels communs aux universités et organismes de recherche (unités, thématiques, auteurs...).

C'est un projet multi-partenaires impliquant universités, organismes, opérateurs et le ministère qui vise à construire et assurer la pérennité d'un référentiel bibliographique fiable fournissant une richesse d'informations maximale pour disposer d'une vision la plus complète possible de la production, éviter de saisir x fois les mêmes données sur les publications, valoriser les activités des professionnels de l'IST...

La méthode de construction privilégie le pragmatisme : utiliser les ressources dont on dispose puis s'organiser progressivement pour minimiser les efforts de chacun.

Il ne s'agit pas de construire une archive, un portail, un outil décisionnel ou un système d'information recherche mais de mettre en place un référentiel alimentant et se nourrissant des dispositifs existants.

L'expérimentation qui a été menée a permis de démontrer la « faisabilité » et l'apport d'un tel dispositif :

- Eviter d'avoir à saisir ce qui l'est déjà ailleurs grâce à un « entrepôt commun de notices descriptives »
- Contribuer à l'amélioration et à la mise en place de référentiels communs (structures, tutelles, auteurs, colloques...).

Commentaires

- Il est clair que disposer des données des bases bibliographiques internationales telles que le WoS sera des plus utiles.
- Il est prévu de réaliser le projet sur 2 ans en impliquant les partenaires dans toutes les phases du projet et en se partageant les tâches et responsabilités. Un financement complémentaire est cependant nécessaire.

Conditor, vers un référentiel national des publications scientifiques françaises

Une recommandation de BSN3 (signalement)
Construire un référentiel de la production scientifique française en s'appuyant sur des référentiels communs aux universités et organismes de recherche (unités, thématiques, auteurs,...)

Colloque Edition scientifique (Meudon)

Atelier 3 : Quelles dynamiques pour valoriser l'information et la publication scientifique ?

4 novembre 2014

BSN3/Conditor_2014-11-04

Synthèse de l'atelier 3

- Les offreurs privés avancent vite et bien avec une démarche « Recherche » et une logique commerciale conduisant parfois à susciter des besoins.
- Les académiques sont bien « réveillés » et ont une pleine conscience de l'évolution des pratiques et des besoins mais ont peut-être un peu de « retard » sur la façon de faire des outils.
- Les objectifs 2015 pourraient être de :
 - Travailler sur une culture du partage des données
 - Monter un workshop technique sur les liens entre publications et données en présence des offreurs de service privés et des académiques.

Conférence-Débat

Maitre Alain BENSOUSSAN, Avocat à la Cour

Droit numérique et plateformes de science : avancées 2014, projets 2015

Les enjeux sont le libre accès et le partage universel de l'IST car toutes les organisations ont, aujourd'hui, un système d'information scientifique et technique.

Il faut trouver un équilibre avec les éditeurs, les participants et les chercheurs en tenant également compte des subventions publiques.

La tendance est l'open science qui doit nécessairement trouver un cadre juridique qui n'est pas limité à la protection par le droit d'auteur.

La réglementation existante est éparse :

- le droit de la recherche
Le principe est l'accroissement des connaissances et le partage de la connaissance
- le droit des données
Aujourd'hui tout est fait pour que ces dernières ne soient pas libres. En effet le droit d'auteur est appliqué à ces dernières et limite.
- le droit du numérique
Ce droit est composé traditionnellement du droit des systèmes d'information, du droit applicable à la protection de la vie privée et du droit de la propriété intellectuelle.
Mais de nouvelles questions apparaissent aujourd'hui : qui va être propriétaire de ces données ? Le financeur ? Le chercheur ? L'auteur ?
Comment peut-on maîtriser l'information ?
Quel est le droit applicable aux contenus générés par les utilisateurs ?
- le droit des plateformes
Il a tout d'abord fallu définir le terme « plateforme » qui est, d'après lui, une concentration de données homogènes à un instant T. Il n'existe aujourd'hui aucun cadre juridique.

Plusieurs rapports traitant de la question des plateformes ont été rédigés durant l'année 2014. Tout d'abord celui de l'Académie des Sciences qui prône l'Open Access Institutionnel. Le Conseil d'Etat a, de son côté, fait cinquante propositions dans son rapport « Le numérique et les

droits fondamentaux », il a rappelé au gouvernement la nécessité de légiférer sur le droit des plateformes et il s'est penché sur la définition ainsi que sur les principes et engagements applicables à l'utilisation de ces dernières.

Enfin le Conseil National du Numérique recommande de favoriser l'interopérabilité et la portabilité des données ainsi que le développement d'un modèle numérique ouvert. Il a également pris position dans son rapport « la neutralité des plateformes ».

L'agenda proposé pour 2015 est la mise en place de groupes de travail avec une participation massive de l'ensemble des acteurs, le recensement de tous les usages, la rédaction d'un livre blanc pour délier le droit de la science du droit d'auteur. Ce livre blanc aura pour objectif de recenser l'ensemble des usages par tous les acteurs du monde scientifique, et notamment les pouvoirs publics, les chercheurs, les éditeurs, l'ADBU et Couperin afin de rédiger un texte consensuel qui aboutirait à de propositions pour la Science. Il pourrait servir d'instrument de discussion tant au niveau national que communautaire.

Présentation

Alain Bensoussan Avocats
Le droit du numérique et des technologies avancées

**Droit numérique et plateformes de Science :
avancées 2014, projets 2015**

*Un Agenda 2015 pour la science publique ouverte
Meudon - 4 et 5 Novembre 2014*

cnrs

Alain Bensoussan
4 novembre 2014

03/12/2014 Copyright Lexing 2014 © Confidentiel Entreprise 1

Un grand projet en 2015 : les données publiques et la science

Henri VERDIER

Administrateur général des données, Services du 1^{er} Ministre

Henri Verdier a reçu une formation par la recherche puis s'est consacré à l'entreprise privée tout en maintenant des attaches dans le Public où il a passé 10 ans dans l'Education Nationale. La formation par la recherche a d'ailleurs été excellente pour construire l'expérience d'entrepreneuriat car on y apprend à faire face aux problèmes en cherchant les solutions déjà existantes et on construisant des équipes d'experts.

Henri Verdier a été nommé, il y a deux ans, directeur du service du premier ministre qui s'occupe de l'ouverture et du partage des données publiques : ETALAB.

Depuis 6 semaines l'Etat construit une fonction de chief data officer en français Administrateur Général des données.

Ces deux services du premier ministre n'ont pas d'autorité directe sur les services de recherche, ni sur le CNRS ni sur les Universités. Il s'agit plutôt d'échanger les concepts et les bonnes pratiques et de se renforcer mutuellement.

L'administration est héritière de concepts qui sont nés dans les campus : les stratégies d'ouverture, d'open source, d'open data, d'open innovation.

Pour l'agenda 2015, ETALAB a déjà pris avec le CNRS quelques rendez-vous. Les deux missions qu'il porte sont jointives et il y a une belle synergie des équipes.

L'Open data est une réalité ancrée dans une réalité française qui remonte à la déclaration des droits de l'homme. Aujourd'hui, ces pratiques sont télescopées par la révolution numérique du web2.0. On assiste à la puissance de la multitude, à la capacité qu'ont les humains de s'approprier les données, de créer de la valeur.

Cette politique d'Open data travaille à la croisée de deux dimensions, **d'une part le socle juridique** (la société a le droit de demander des comptes à tout agent public et son administration – déclaration des droits de l'homme) de ce point de vue, l'organisme public qu'est le CNRS doit rendre des comptes à la société. D'autre part, il y a une très longue et patiente

construction de la transparence démocratique, il y a une autorité de contrôle faite de magistrats indépendants qui critiquent l'état dans un rapport, lequel est public depuis 1850 – Cour des Comptes. Lucien Bonaparte a construit un service de la statistique publique et en 1870 sont nés « les bulletins de la statistique publique », ancêtres de l'INSEE. Puis la France a construit le service public de l'information statistique en 1951 et la loi CADA (Commission d'accès aux documents administratifs) a été votée ce qui fait que tout citoyen français peut réclamer la transmission de tout document administratif, tout en incluant un certain nombre de barrières de sécurité protégeant la vie privée de ceux qui sont concernés.

Une tradition liée à l'ouverture des données, qui est née à Stanford, Harvard ou le MIT, et qui est de créer des zones de coopération, des alliances, de révéler la valeur.

Cette évolution ne vient pas de la révolution française mais des pratiques de la science et du numérique. Internet c'est d'abord un projet politique qui vient des chercheurs, aujourd'hui il est en train de changer pour devenir un projet commerçant.

A la croisée de ces deux grandes vagues d'innovation, dans l'état actuel de la réflexion du gouvernement, on favorise l'ouverture la plus vaste, la plus gratuite, la plus ré-appropriable possible des données publiques tout en étant capable d'y inclure des paramètres de sécurité. Un parallèle grossier peut être fait entre ETALAB et l'organisation du commerce. En effet, si tout le monde partage ses savoirs, le système peut fonctionner mais si vous êtes le seul à partager avec un partage unilatéral, vos savoirs seront absorbés par toutes les autres plateformes et vous serez mis en danger. ETALAB veille à construire une libéralisation progressive de tous ces savoirs.

Le grand débat public confié au CNNum portant sur une loi sur le numérique a été cité hier. Probablement les projets de législation seront testés dans ces débats dans les 4 mois qui viennent, on s'en servira pour co-écrire la loi numérique en temps réel.

Un des quatre chapitres de la loi concerne la modernisation de la fonction publique dans lequel il y a un chapitre open data et un chapitre open-gouvernement.

Il faut y participer, ne pas hésiter à y aller, à poser des contributions. Ce sera vraiment utilisé car on cherche encore des points d'équilibre.

Une des complexités de l'open data et de la réflexion que nous menons est que l'on mélange trois objectifs :

On travaille cette ouverture des données publiques à des fins de **démocratie** plus aboutie. La **transparence** est un fondement de la démocratie mais elle ne suffit pas car elle doit être liée à **la responsabilité de rendre des comptes**. Or la transparence passive, nue, n'est pas le reflet de l'exercice de notre responsabilité.

La libération des données a lieu aussi car elles sont le ferment parfois, de beaucoup d'innovation et de valeur. Il y a une centaine de start-up qui utilisent les données publiques. Mais il y a aussi de grands industriels qui les utilisent. Les infrastructures critiques de la modernité seront des infrastructures de données. Un exemple que l'on connaît tous, le GPS, service mobile qui existe grâce à une infrastructure de données et qui est contrôlé par ceux qui détiennent cette infrastructure de données.

Cet exemple pourra être reproduit sur des domaines aussi variés que l'énergie, le transport, la santé. Il faudra trouver les dosages subtils, entre une attitude fermée, et une logique d'abandon des données, d'abondance et d'ouverture totale. Il faudra comprendre intimement comment on injecte de la valeur dans l'économie, dans la société en construisant des modèles de données.

Pour comprendre à quoi peut servir l'Open data, on peut imaginer un monde sans cadastre, sans données IGN, météo, et se demander que deviendrait l'économie française ?

L'Open data, c'est donc une démocratie plus aboutie, plus interactive, elle permet la création de valeur et constitue un ressort de l'efficacité de l'action publique.

Il existe malheureusement encore des cloisons entre les services de l'Etat. Il y a des cloisons entre les données. Par exemple, la Ministre du logement, ne connaît pas les prix de l'immobilier et les achète aux Notaires, à la fin de l'année écoulée pour 100 000€/an. Elle reçoit ainsi une photo statique de l'année écoulée. Il est difficile de piloter une politique du logement sans pouvoir obtenir des retours rapides et interactifs, du type Test & Learn.

Il y a également, l'exemple de deux ministères différents qui mesurent la production des nappes phréatiques et la pollution récurrente et qui ne croisent pas leurs modèles. De même, les douanes et l'administration fiscale ne croisent pas leurs données, non plus. Ces cloisonnements induisent une perte de valeur importante. De plus, la non exposition de ces données empêche leur critique, leur appropriation et par conséquent, leur amélioration.

Un important travail sera effectué autour de la Base Adresse Nationale (BAN). Depuis déjà 12 ans, le conseil national de l'information géographique fait remarquer à quel point il serait indispensable que chacun puisse bénéficier d'une base nationale de géolocalisation exacte avec toutes les adresses en France. Elle bénéficierait aux Samu, aux pompiers, aux opérateurs télécom, à la poste, pour qui connaître la géolocalisation exacte de la personne à visiter peut s'avérer essentiel.

Il y a ainsi une demi-douzaine de bases en France : toutes incomplètes, fermées, fausses et incapables de se debugger entre elles. D'un autre côté, une application appelée OpenStreetMap (sorte de Wikipedia de la carte), qui bénéficie de 10 000 contributeurs. ETALAB a proposé son aide à OpenStreetMap pour créer une base nationale ouverte.

Pour cela, ETALAB a construit la petite brique (une année ingénieur de développement) qui permet de gérer un modèle recevant des flux de données de sources très différentes. En effet, il était attendu que l'IGN et les cadastres interviennent d'une façon ou d'une autre et il s'agissait de permettre tous les business models à venir.

Cet accord sera signé la semaine prochaine. Il réunira l'Institut National de l'Information Géographique et Forestière (IGN), La Poste, l'Administrateur Général des Données et la mission ETALAB du Secrétariat Général pour la Modernisation de l'Action Publique (SGMAP), ainsi que l'association OpenStreetMap France. Ils s'associent pour constituer ensemble une Base Adresses Nationale (BAN), <http://www.etalab.gouv.fr/acteurs-publics-et-societe-civile-sassocient-pour-la-constitution-dune-base-adresse-nationale-ban-collaborative>. Ils coopèrent ensemble pour créer un modèle de données qui à ce stade est non attribué et mélange toutes les sources d'information.

OpenStreetMap mettra à disposition ce résultat gratuitement, en open data avec une licence ouverte réutilisable à l'identique, car le pouvoir est chez celui qui arrive à mettre à jour ses données, le plus vite possible.

En effet, cela évitera de renouveler l'expérience de l'IGN qui a vendu pour 14 millions d'€ à Google les fonds de carte de France. Ce dernier a décliné l'année suivante l'offre des mises à jour, car sa connaissance des déplacements en temps réels lui permet de connaître les évolutions des sens uniques et des limitations de vitesses sur les routes.

Le pouvoir se trouve bien chez celui qui peut mettre à jour et non pas chez celui qui possède la base de données initiale. D'où l'utilisation d'une licence où le partage est gratuit et où la restitution du bien commun à la base initiale est prévue. Dans ce cas l'IGN conserve ses pratiques actuelles pour ceux qui souhaitent une information certifiée par une autorité. C'est un **modèle de double licence**, qui est mis en place pour la première fois.

Il y a peu de culture de data driven stratégie dans l'administration. Les décisions sont prises sans tentatives d'étudier à l'avance les effets de ces décisions. Il faut injecter de la stratégie dans la conduite des affaires publiques, avoir le réflexe de revenir aux faits, de voir si on a des données pertinentes, de construire des modèles de données qui aident à prendre la bonne décision, qui aident à piloter les politiques.

Il faudra inventer les stratégies fondées sur la donnée, les tester, les théoriser etc.

Il faut aussi revoir les stratégies d'achat : un manager « à l'ancienne » pourrait se dire qu'il suffit de connaître les mauvais acheteurs, pour les licencier. On peut choisir cette méthode, mais on peut choisir aussi de construire un modèle d'évaluation afin que chaque acteur connaisse les meilleures pratiques de ces collègues et que tous ensemble s'entre-améliorent. Savoir expliquer,

théoriser, et implémenter ce genre de stratégies, et la variété des stratégies fondées sur les données est une chose qui manque en France en général.

L'Etat fournit des statistiques issues de ses données. Les statistiques sont nécessaires, néanmoins, l'existence d'un référentiel sur lequel les personnes intéressées peuvent se synchroniser est essentielle. En effet, la simple attribution d'une donnée pertinente à celui qui prend une décision sur le terrain est une aide à la décision très précieuse et correspond à la boucle de rétro-control de la méthode du test & learn.

Pour conclure, deux missions reviennent à Henri Verdier :

L'ouverture des données, de manière à ce qu'elles soient réutilisables, afin que quiconque puisse les retravailler, les tester. Il ne suffira plus que les « sachants » énoncent des résultats, le citoyen pourra les vérifier.

Les données seront ouvertes le plus possible, gratuites le plus possible et le plus réutilisables possible sans restriction, ni partage à l'identique. Il faudra travailler sur les conditions de licence de partage à l'identique.

Travailler sur des modèles de revenus : c'est la théorie des communs :

Quand on a un bien commun, il faut que ceux qui en tirent le plus de profits contribuent à la maintenance du commun, sinon il risque de ne plus y avoir de commun.

La valeur de la donnée vient de ceux qui l'utilisent pour innover et donc ceux qui lui trouvent des usages. ETALAB est favorable aux modèles dont le budget est proportionnel à l'utilisation ou à la valorisation ou au succès, plutôt qu'aux modèles qui mettent une barrière pour l'accès dès l'entrée et qui décroissent en volume avec le temps.

La mission de l'administrateur général des données est de faire en sorte que les décisions soient davantage fondées sur la donnée.

On rencontre trois familles de problèmes autour de :

- La circulation de la donnée (interopérabilité, coût,...)
- L'apprentissage de l'Etat à l'utilisation de la donnée, donc s'organiser pour devenir un hub et favoriser des collaborations, les créations de start-ups. L'Etat attend beaucoup du CNRS et des autres établissements sur ce sujet.
- La diffusion de la culture de la donnée

Henri Verdier a déjà pris date avec le CNRS pour étudier parmi les données produites par la recherche, celles qui permettent d'évaluer des politiques publiques et pour cela des expérimentations seront menées. De même, le CNRS sera sollicité pour la conception de modèles pour les données de santé, d'achat public, d'emplois, d'aides aux entreprises (le lab de

la BPI (<http://www.bpifrance-lelab.fr/>) possède par exemple 100 ans d'archives de la caisse des dépôts).

Puis dans les réflexions à venir sur l'efficacité des modèles d'ouverture de données, et sur la durabilité des business models.

ETALAB sera heureux de partager les réflexions lors du séminaire organisé par le CNRS en février 2015.

La recherche a fait la preuve de son efficacité grâce à des modèles fondés sur le jugement par les pairs (peer review) la vérification, la fiabilité, l'exposition et l'acceptation du risque de la critique. Ce sont ces valeurs qu'ETALAB espère mettre en œuvre dans d'autres pans de l'administration publique car sans être chercheur on peut être responsable d'une politique de santé, de logement, et adopter des valeurs similaires.

C'est dans le monde de la recherche que sont nés les concepts d'open data, open innovation, openscience. Merci de continuer à accompagner les autres organisations.

L'Information scientifique en 2015 : les facteurs du changement

Les besoins : résultats de l'enquête auprès des Directeurs d'unité de recherche

Francis ANDRE (DIST- CNRS)

Dans le cadre de la mise en œuvre de la stratégie IST du CNRS, la DIST a mené cet été deux enquêtes en ligne, l'une à destination des directeurs d'unité, l'autre à destination des personnels d'appui à la recherche dans les fonctions IST.

Le dépouillement de ces deux enquêtes doit permettre au CNRS de disposer d'une vue précise des pratiques et besoins d'IST dans les unités du CNRS, dans les différents domaines couverts par les Plans d'Actions Partagées (PAPs) : documentation, publication, valorisation de l'information et accompagnement de l'IST. Les conclusions attendues en début d'année 2015 permettront de valider avec les instituts de nouvelles orientations des services IST.

Volet Directeurs d'unités :

432 unités ont répondu aux 91 questions de l'enquête ; cet échantillon est parfaitement représentatif de la diversité du CNRS. Cette diversité se traduit par des comportements et avis des instituts très contrastés sur des sujets tels que l'usage et l'obligation de dépôt dans l'archive

ouverte HAL, le regroupement des portails d'accès aux revues électroniques, les pratiques de gestion des données produites par les laboratoires... Ceci conduit d'ores et déjà à envisager des actions ciblées par grands domaines scientifiques à définir précisément avec les instituts.

Sans grande surprise apparaît nettement, dans les champs disciplinaires concernés, la nécessité d'investir dans la valorisation des données de la recherche, tant au niveau des moyens à y consacrer que des compétences à développer.

Volet professionnels de l'IST :

265 réponses à 162 questions, là encore un échantillon représentatif de la diversité des pratiques des instituts. Les premières analyses font apparaître une évolution des activités des professionnels IST dans les laboratoires : par exemple, aux traditionnelles fonctions d'aide à la recherche d'informations se substituent les fonctions de valorisation des productions des chercheurs (publications dans les archives ouvertes, revues électroniques, métadonnées), entraînant *de facto* l'expression de besoins de formations.

L'innovation éditoriale

Olivier DUMON (Elsevier)

Mondane MARCHAND et Nicolas NOQUET (Thomson-Reuters)

Claude KIRCHNER (Inria)

En introduction de la session, Claude Kirchner décrit la qualification et la certification des documents qui doit être séparée de la publication quel qu'en soit son mode, pages web, archives ouvertes, blogs... En quelque sorte, la séparation des pouvoirs permet une démarche plus ouverte.

L'un des enjeux majeurs concerné les données, *i.e.* les résultats de la recherche scientifique qui doivent être rendus publiques, ce qui est rendu possible par les archives ouvertes.

L'innovation éditoriale : les objets sont déjà rendus publics suivant les communautés et les besoins, les attitudes éditoriales sont toutefois différentes. Des innovations peuvent être expérimentées.

Un comité éditorial avec des referees décide de l'acceptation du papier tout en générant des données d'évaluation qui sont alors sous la maîtrise de la plateforme publique ou privée. Ce processus standard a des variantes, quelquefois les rapports des referees et même leur identité sont rendus publics, c'est ce qui se fait dans le domaine de l'anthropologie notamment. En tant qu'auteur, il est nécessaire de se rendre compte qu'en soumettant un papier via des plateformes

privées, on perd la maîtrise des données produites et en conséquence le partage de toutes ces données.

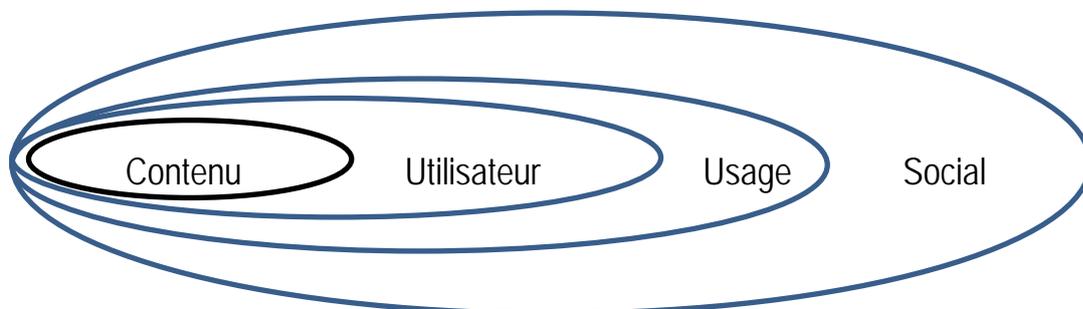
[Episciences](#) est une alternative qui permet de développer des comités éditoriaux en gardant la maîtrise des données dans un cadre public : la première étape est le dépôt dans HAL ou ArXiv suivi ensuite d'une certification via [Episciences](#). En mathématiques, Jean-Pierre Demailly met en place un [méta comité éditorial](#) chargé de sélectionner les journaux qui vont être mis en place dans ce cadre. Le comité éditorial et les communautés de chercheurs peuvent développer leurs idées sur ce nouveau réseau social ainsi constitué. En informatique et en mathématiques appliquées, deux revues existent d'ores et déjà.

La notion de plateforme éditoriale peut être généralisée.

Le futur de [Hal](#) : par le dépôt dans HAL, l'objet est ensuite soumis à une communauté scientifique dont le comité scientifique. La qualification est alors composite, par le comité éditorial mais aussi la communauté de chercheurs en parallèle.

Elsevier a ainsi bien cerné cette évolution par le rachat de Mendeley, logiciel de gestion bibliographique, destiné à la gestion et au partage de travaux de recherche. Il agglomère ainsi les données et rend possible de générer des liens entre l'objet, les données et l'évaluation via ce réseau social.

Pour Olivier Dumont, Elsevier, l'innovation éditoriale s'inscrit dans le contexte du besoin du chercheur qui est en évolution à toutes les étapes. La valeur éditoriale apportée par les éditeurs doit changer. L'internet basé sur les moteurs de recherche passe aussi par un internet de réseau. Selon un des fondateurs de LinkedIn, la production de contenu était ce qui définissait les acteurs de l'internet. Maintenant c'est le réseau qui prend le dessus, la façon dont il est relié à l'écosystème. La publication scientifique reste bien entendu au centre de l'information avec 3 niveaux d'information à mettre à disposition des chercheurs :



L'écosystème d'Elsevier est constitué de 7 plateformes (EEE, Evisé - plateformes de soumission d'articles, ScienceDirect, SCOPUS et Mendeley qui sont multi provenances, SciVal et PURE) à

mettre en réseau via des services d'API. L'intégration des plateformes est déjà effective avec [ORCID](#), [CrossRef](#) et 40 dépôts de données de recherche.

Sous-jacentes à tout cela, les questions de vie privée sont un point sensible. Le partage des informations ne se fera qu'avec l'accord des chercheurs. Un équilibre est à trouver entre partage et respect de la vie privée.

Elsevier est demandeur de partenariat sur le thème de la recommandation (Mendeley), sur l'intégration avec HAL (apporter par exemple des profils utilisateurs SCOPUS pour améliorer les métadonnées), sur l'amélioration des adresses.

Pour Nicolas Noquet, Thomson-Reuters, les métiers sont en profonde mutation pour répondre à des besoins spécifiques. Thomson, est un industriel de l'information, éditeur de solutions reconnues mondialement. Le Web of Science a profondément évolué et des formations sont gratuitement disponibles bien que sous-utilisées. Des adaptations régionales sont faites par rapport aux innovations mondiales de l'éditeur. Voir notamment le partenariat avec l'outil SAMPRA développé par le CHR de Lille 2, qui permet la validation au niveau chercheur des publications et le rattachement possible de ces informations au niveau des structures, départements et institutions. Les informations connectées consolidées permettent ensuite une visibilité accrue au niveau international. Cette solution est gratuite pour tout abonné JCR/WOS pour améliorer son repérage et sa visibilité interne ou externe.

Mondane Marchand, Thomson-Reuters, décrit des innovations éditoriales de Thomson-Reuters et notamment le partenariat établi avec Google Scholar, qui est gratuit pour les abonnés WOS et permet aux utilisateurs de passer d'une plateforme à l'autre après activation des liens. Thomson-Reuters met à disposition 3000 revues en open access dont le full text est donc accessible et a signé des accords avec les académies des sciences en Chine, Corée et est en discussion avec la Russie et l'Inde. L'objectif est de valoriser les collaborations et la littérature en open access notamment en provenance des pays émergents.

Thomson Reuters a été le premier à lancer une plateforme permettant aux chercheurs de répertorier l'ensemble de leurs publications via les Research ID, technologie aujourd'hui utilisée par ORCID et travaille maintenant au développement de nouveaux indicateurs.

Nicolas Noquet annonce la volonté de Thomson Reuters de proposer des accès à l'ensemble de ses index pour les PRES et COMUE en France.

Les plateformes multi-usages : pratiques actuelles

Christine BERTHAUD (CCSD- CNRS)

Anne CAMBON-THOMSEN (CNRS)

Renaud FABRE (DIST-CNRS)

Anne Cambon Thomsen, médecin et directrice de recherche au CNRS travaille dans le domaine de la santé.

En biologie, il existe beaucoup d'infrastructures qui se développent et c'est un domaine où la donnée n'est pas uniquement une donnée qui peut être numérisée, il faut également tenir compte de la source des données, que sont les échantillons et collections d'échantillons biologiques. Ces collections d'échantillons n'ont aucune valeur si les données descriptives n'y sont pas associées.

Dans la production scientifique aujourd'hui, outre les articles, il existe la constitution de bases de données de collections d'échantillons et qui, lorsqu'il s'agit d'échantillons humains, sont d'autant plus sensibles.

Cette constitution d'échantillons de bio-ressources n'est pas considérée, reconnue, comme la production d'une publication.

On constate que ces bio-ressources sont une source extraordinaire, et de plus en plus importante de production de résultats scientifiques.

Or c'est un travail important de rassembler ces échantillons physiques et d'organiser les données associées dans un travail de qualité. Et finalement, on trouve une résistance à ce partage car si on donne accès à la base, on perd le contrôle et donc le contrôle de la qualité.

A. Cambon participe à des groupes de travail nationaux et internationaux où est traitée la question de la manière de procéder pour valoriser ce travail de constitution des ressources et de le rendre partageable, car la mise en ligne, seule, n'implique pas le partage.

C'est une réflexion globale qui intègre la création d'index tel que le bioresource research impact factor (<https://www.bioshare.eu/content/bioresource-impact-factor>)

Comment suivre l'utilisation de ces ressources et leur attribuer à travers ces mesures, un impact sur la production scientifique.

Pour ce travail il faut pouvoir :

- Identifier les ressources (pas d'identifications uniques pour ces ressources) et les data sets, ainsi que les ressources physiques qui sont à l'origine des données.
- Tracer l'utilisation (mesure de l'utilité pour la science) et plus c'est utilisé, plus c'est cité et plus l'index augmente et donc intérêt à partager.

L'idée est de pouvoir avoir dans le profil ORCID la documentation indiquant la contribution à telle bio-ressource ou à tel identifiant. Ainsi on attribue une reconnaissance d'avoir contribué à une bio-ressource dont on sait mesurer l'utilité car il existe un identifiant et on pourrait avoir des standards de citation.

A. Cambon a un article soumis sur la proposition de standard d'utilisation de bio-ressources utilisé par des articles. Il s'agit également de faire la différence entre une bio-ressource qui a contribué au résultat d'un article et une bio-ressource qui est citée dans un article.

Un travail important est mené avec les éditeurs sur des standards de citations décrits actuellement sur le site « equator ». C'est ensuite aux éditeurs ou aux reviewers de vérifier que ces standards soient appliqués.

La qualité et le partage de ces données et sources de données reposent donc en grande partie sur la capacité à pouvoir mesurer leur utilisation et reconnaître le travail qui est à leur origine.

Ce travail se place dans le contexte général des plateformes des sciences de la vie qui ont une certaine ancienneté et qui se situent autour d'appareillages qui possèdent des chartres d'utilisation, et ont déjà une histoire dans la construction de plateformes technologiques.

Ces plateformes aboutissent à la création de ressources qui peuvent être des échantillons, des logiciels, des outils mis à disposition.

Ibisa est l'infrastructure nationale de coordination des plateformes dans le domaine des sciences du vivant depuis 20 ans. Elle a mis en place des chartres de fonctionnement, des requis d'ouverture pour l'utilisation de ces plateformes qui doivent être ouvertes au public. Il existe également des réseaux de plateformes dans le domaine de la bio-informatique, de la chimie, de l'imagerie. A l'origine, seule la génomique était concernée, mais aujourd'hui, une chartre commune des plateformes est mise en place.

Un groupe de 135 personnes environ travaille sur cette notion de ressource, et de reconnaissance de l'impact de ces bio-ressources en sciences et il est en train d'intégrer l'infrastructure européenne BBMRI biobanking and bioMolecular resource infrastructure <http://bbmri-eric.eu/> qui a le statut d'ERIC (European Research Infrastructure Consortium http://ec.europa.eu/research/infrastructures/index_en.cfm?pg=eric). BBMRI s'intéresse dans le domaine humain à toutes les bio-banques, collections d'échantillons et données associées et organise, au niveau européen, avec des non-nationaux le réseau de ces banques. Ces bio-banques ont un identifiant, un catalogue en cours de constitution. Le concept est très général et peut être appliqué à des collections en dehors du domaine des biologies humaines (musées, etc.).

L'Inist souhaiterait avoir la possibilité d'étudier cette perspective sur place.

Présentation

bbMRI-lpc **BioSHaRE.eu** **GEN2PHEN** **Gage Kic**

BRIF
Bioresource Research Impact Factor

**The Bioresource
Research Impact Factor
initiative**

Anne Cambon-Thomsen (*leader*), Laurence Mabile (*manager*)
UMR1027 Inserm-Université Toulouse III, FR

Paola De Castro, Federica Napolitani, Anna Maria Rossi, Alessia Calzolari
and Elena Bravo (*BRIF & Journal Editors' subgroup*)
Istituto Superiore di Sanità, Rome, IT

Inserm **UNIVERSITÉ TOULOUSE III PAUL SABATIER**

HAL

HAL se situe au cœur d'un réseau qui :

- Développe des plateformes en interne du CCSD, telles que sciencesconf qui est une plateforme d'organisation de conférence, et episciences plateforme développée depuis deux ans et demi grâce au soutien de l'institut Fourier, JP Demailly, et Inria qui a démarré deux revues sur cette plateforme
- Est interconnecté avec ArXiv (Hal se positionne ainsi en tant qu'Archive internationale)
- Travaille avec le CINES - réplique de la base au CINES et gestion de l'archivage pérenne, sauvegarde du support des documents pour y accéder dans des dizaines d'années
- Est partenaire ANR AR campus avec un projet de diffusion d'archives ouvertes, de vidéo
- Est partenaire de DARIAH (projet européen)
- Assure la Maitrise d'œuvre du moteur de recherche Isidore permettant au CCSD de monter en compétence
- Est localisé dans les locaux du Centre de calcul de l'In2P3 ce qui lui permet de bénéficier d'un réseau performant.

La V3 est en ligne après la réalisation d'un travail important fait sur les tests. C. Berthaud explique qu'elle est encore en période de « débogage » mais le travail avance vite.

HAL sert des communautés disciplinaires et institutionnelles très différentes. Dans le cadre de la convention inter-établissements, HAL propose des portails institutionnels pour les institutions qui souhaitent créer leurs archives au sein de HAL et HAL sert d'archives ouvertes communes et

partagées. Dans le cas où l'archive institutionnelle est gérée localement, HAL récupère les données et assure l'archivage pérenne.

HAL contient des archives disciplinaires variées tel que celle des sciences humaines et sociales qui fêtera ses 15 ans en 2015 ou celles liées aux typologies des documents comme l'archive ouverte de Thèses.

HAL est fondée sur le document qui est au cœur de l'archive ouverte.

HAL s'appuie donc sur des référentiels pour la gestion de ses documents :

- Référentiel des laboratoires et des institutions (et référentiel auteur) qui gère les affiliations.
- Référentiels des domaines scientifiques, celui des typologies de documents (basés sur le référentiel du ministère, élargi à une typologie de type vidéo).

Les archives institutionnelles accueillies dans HAL peuvent gérer et personnaliser leurs pages web, gérer leurs chartes graphiques, des news. De nombreux services sont ainsi offerts par HAL, leur permettant de gérer leur archive comme si elle était hébergée dans leur établissement, d'administrer des collections pour mettre en avant des projets de recherche, des laboratoires.

HAL peut également récupérer et traiter des statistiques de consultations et d'usage.

HAL permet donc de garantir l'accessibilité au texte intégral, de s'interconnecter aux grandes archives internationales, de préserver un niveau scientifique homogène et de qualité, et de limiter la saisie de références à un seul système.

Services à forte valeur ajoutée mis en place dans HAL :

- Gestion des listes de publications (exportation et exploitation libres)
- Alimentation des bases administratives (bilans annuels des chercheurs)
- Partage du dépôt avec tous les coauteurs quels que soient leur établissements en France ou à l'étranger
- Création d'épi-journaux
- Dépôt de données de recherche, de vidéos
- Liaison de ressources entre elles
- Lieu de stockage unique situé en dehors des circuits privés
- Point d'entrée pour le moissonnage de tous types de plateformes (institutionnelles, réseaux sociaux).

Identifiant unique

Création d'IdHAL : l'identifiant du chercheur, qui permet de générer et choisir des données qu'il souhaite utiliser ou pas. Depuis 2 semaines (lancement de la v3) 856 identifiants ont été créés. Il y a eu 366 pages scellées. Cet identifiant permet de lier d'autres identifiants.

Parmi ces 856 idHAL, 499 IdHAL sont liés à d'autres systèmes dont 102 researcher ID, 105 Orcid, 96 arXiv, 57 linked in, 41 Idref. Cela indique bien que le chercheur s'intéresse aux évolutions de HAL.

Dariah :

Infrastructure de recherche liée aux sciences sociales et humaines dont l'objectif est de construire une communauté européenne numérique, en mettant en place des bonnes pratiques, etc. Le CCSD offre la mise en place d'un portail spécifique et l'outil scienceconf d'aide à l'organisation des événements liés à Dariah.

Dans HAL l'auteur choisit sa propre licence creative common.

Conclusion R. FABRE

Dans ces 2 présentations on a vu apparaître un rôle nouveau dans le travail de recherche qui est d'inventer en permanence les catégories à partir desquelles on peut utiliser les moyens numériques pour développer le travail de la science.

Les choses n'ont pas encore toutes un nom. Certaines s'appellent nanopublications, bibliothèques sémantiques...

Tous ces objets, sont en cours de construction. Travailler sur l'agenda 2015, permettra d'organiser l'effort, de le partager et de chercher ces nouvelles catégories d'objets.

Par exemple, l'atelier 3 sur les plateformes qui s'est tenu autour de S. Pouyllau a fait des propositions pour développer des outils à partager entre les plateformes.

L'utilisation de la plateforme CorIST a permis de développer la stratégie IST du CNRS en proposant à chaque communauté de s'emparer de l'outil, des documents, de partager des informations sur les disciplines, de réagir, et de couvrir des territoires beaucoup plus vastes que ceux utilisés habituellement.

Cela a permis de construire les 4 plans d'actions partagés du CNRS et la cinquantaine d'actions touchant au domaine de la publication, de la documentation, de l'analyse ou de l'accompagnement de l'IST. La plateforme CorIST a conduit à, l'augmentation de la fréquentation du site de la DIST (passée de quelques centaines à quelques 4400 consultations mensuelles) et à créer une confiance avec la communauté des chercheurs qui a abouti aux deux enquêtes présentées le matin.

Il conclut en donnant une esquisse d'une possible typologie des plateformes :

- **Plateformes de recherche** qui contiennent des données, des analyses mais aussi des équipements de recherche. Elles collaborent à la production de la science, comme au génopole
- **Plateformes d'analyse** dans lesquelles on trouverait des données, des publications et des analyses. Dans ces plateformes, on chercherait à produire du sens sans avoir accès à l'expérimentation
- **Plateformes d'exploration** dans lesquelles outre des données, des analyses et des articles, on trouverait aussi la restitution des contenus générés par les utilisateurs, soit l'ensemble des démarches par lesquelles une communauté progresse dans son travail.

L'Information scientifique en 2015 : vers un agenda partagé

Débat animé par

Jean-François Balaudé, Président de l'Université Paris X Nanterre

Didier Torny, Directeur scientifique adjoint InSHS, CNRS

Avec

Ghislaine Chartron, groupement français des industriels de l'information (GFII)

Grégory Colcanap, coordinateur du consortium Couperin

François Geze, Président du groupe des Editeurs en SHS du syndical national de l'édition

Emmanuelle Jannes-Ober, présidente d'Eprist

Joanna Janik, DIST-CNRS

Christophe Péralès, Président de l'ADBU

La question, rappelée dans le rapport ADBU-Eprist, de l'articulation entre politique nationale et politique de site, voire politique européenne et internationale, se pose aujourd'hui. Il est vraisemblable que l'agenda à partager porte sur ces divers niveaux.

Pour Eprist : Il existe un véritable enjeu de clarification de la politique nationale en matière IST, de son existence, de sa coordination. Elle doit être co-construite avec l'ensemble des acteurs. Ce travail collectif de participation à la politique nationale comprend l'ensemble des segments de BSN, incluant les collections d'excellence (collection papiers).

La question qui se pose est : comment alimenter et nourrir une politique nationale ?

Pour l'ADBU : Le constat de la note ADBU-Eprist, est que les questions qui ont trait aux contenus, aux collections, aux données..., doivent se conserver au niveau national. Les sites intervenant au niveau des services de l'accompagnement.

Il y a là donc une politique nationale à construire (BSN, Collex, travail CNRS/Bensoussan/ADBU sur le droit des plateformes, entre autres).

Henri Verdier a bien positionné la question des données et leur ouverture qui constitue un levier économique formidable. Olivier Dumon a rappelé que relier les données entre-elles est un amplificateur non négligeable de cette économie.

Dans la chaîne à mettre en place entre tous les acteurs, y compris les acteurs privés, comment éviter les captations de valeur (comme on a pu le connaître dans les publications) ? Cela implique une coordination au niveau des acteurs publics.

COUPERIN rappelle l'importance du décloisonnement qui a été évoqué par H. Verdier. Couperin propose d'adopter la manière de travailler des chercheurs sous forme de critique constructive et mutuelle et de coopération.

4 champs sur lesquels l'agenda pourrait travailler :

1. **Construire un système d'information partagé** pour faire circuler les données. Des problèmes de référentiels ont été abordés et l'on constate que c'est souvent les difficultés techniques qui rapprochent les communautés. De ces aspects techniques sont nés des projets de politique documentaire.
2. **Construire une politique documentaire** : ISTEEX est un exemple de projet de politique documentaire. Il permet l'acquisition de ressources mais il comporte aussi une dimension recherche & innovation puisqu'il vise à fournir des services innovants au chercheur. Les paiements d'APC et l'atomisation des paiements induisant une dérive totale des coûts sont un exemple de chantier qui doit être instruit très vite, car le temps du numérique diffère de celui du politique. Cet exemple a été pointé par l'Académie des sciences dans son rapport : on ne sait pas combien on dépense. La question de la négociation des APC au niveau national est importante pour la politique documentaire d'un pays et de ses établissements.
3. **Champ juridique** : l'intervention d'Alain Bensoussan a traversé toutes les thématiques abordées. Notamment, la perte de la production des chercheurs quand ils cèdent leurs droits aux éditeurs. Il y a ici toute la question des réseaux sociaux et des données des chercheurs. Il s'agit d'éviter de se déposséder de ses données lors de la signature d'un contrat avec un éditeur.
4. **Chantier politique** : lors du débat « Enjeux d'avenir pour la publication scientifique », il a été fait un état des lieux de l'Open Access en France. Il est temps de passer des diagnostics aux décisions. Il pourrait y avoir une politique d'obligation de dépôt auprès des établissements et ces décisions peuvent aussi se traduire par la voie parlementaire et par des dispositions législatives.

Cet agenda pourrait tracer des perspectives sur ces 4 champs.

S. Pouyllau conclut l'atelier 2 en invitant à mettre en place un **dictionnaire des référentiels**. Ce point décrit comme celui d'un schéma national des données a été abordé à l'ADBU, à la BnF et à l'Abes et cela décrit une convergence très forte. Il faudrait arriver à sérier toutes les données dont on parle, voire les traitements et les chaînes de composition et d'enrichissements dans lesquels elles peuvent s'inscrire.

E. Jannes-Ober rappelle que, dans le cadre de BSN10, il va être proposé de dresser un **état des lieux des données**.

Le SNE appelle de ses vœux une coopération critique, dans le bon sens du terme, entre l'édition et tous les acteurs de l'IST, les bibliothèques et le monde de la recherche. Les modalités de la mise en œuvre de l'Open Access motive le SNE tout autant que les chercheurs. Le SNE souhaite une approche pragmatique et concrète où les dossiers seront discutés un par un. En cela, le rapport de l'Académie des sciences est intéressant car il est pragmatique et prend en compte l'édition scientifique qui est reconnue comme un des maillons fondamentaux de l'IST. Ce rapport reconnaît également que l'édition a un coût.

Selon le SNE, le débat sur le Green et le Gold n'a pas assez été discuté autour d'une table avec tous les acteurs. Le SNE a, par ailleurs, demandé au MENESR la création d'un observatoire du

numérique dans l'enseignement supérieur et la recherche. Un observatoire du numérique a été mis en place dans l'enseignement supérieur (questions du numérique au niveau licence) et a conduit à la constitution de groupes d'études pour comprendre l'usage des ressources numériques dans le monde universitaire.

Mais cela n'a pas été fait pour la recherche où l'usage des ressources numériques est plus ancien, plus structuré et plus important.

Le SNE est très inquiet de **constater l'évolution dramatique des budgets des acquisitions** des bibliothèques universitaires et des centres de recherche, sachant qu'en France ces budgets sont à un niveau 2 à 3 fois inférieurs au reste de l'Europe en ce qui concerne l'acquisition de ressources numériques. Cette diminution est difficile à accepter car tout le reste en dépend, il ne sera plus nécessaire de discuter Green ou Gold s'il n'y a plus de budgets pour financer l'acquisition et la diffusion convenable de l'IST.

Le 2ème point important concerne les embargos qui doivent être différents selon les disciplines.

Il y a des pistes d'entente et il est important d'avancer dans ce sens. **Héloïse** est un exemple de partenariat mis au point par le SNE et le CCSD. C'est un site, à l'instar de Sherpa Roméo, où les éditeurs sont invités à faire connaître leur politique d'Open Access, d'embargo, de preprint, de postprint.

Le dernier point concerne les modalités d'acquisitions des ressources éditoriales numériques pour le monde de l'ESR qui passe par la question des licences nationales à laquelle le SNE adhère mais qui demande à être discutée et pensée conjointement pour se faire au mieux des intérêts de tous.

Le groupe OA du GFII regroupe à la fois des éditeurs, essentiellement francophones, et des EPST. Ce sont donc des acteurs du contenu et du contenant qui essaient de faire le lien entre tous les acteurs impliqués dans cette chaîne de l'IST.

Quel agenda 2015 ? Quelle échéance ? Quel objectif ? L'avantage de cette démarche est qu'elle fait travailler ensemble des acteurs du secteur public et du secteur privé : on ne peut avoir un agenda commun, si on n'a pas créé un commun.

G. Chartron pense que l'on a besoin de tous les talents, y compris ceux du privé, pour proposer un certain nombre de services. Dans ce groupe, il y a eu beaucoup de débats et aujourd'hui quelques fondamentaux réunissent ses membres.

Le modèle de communication n'est pas unique, cela a été largement rappelé lors de ces journées. La transformation numérique ne se fera pas de la même manière selon les secteurs. Les enjeux autour du TDM sont à moduler en fonction des disciplines. L'idée étant de préserver le tissu éditorial national. L'exemple de Scielo en Amérique du Sud, où le modèle d'archives ouvertes a bien fonctionné, le montre car ce sont des pays où les éditeurs sont peu présents. La difficulté pour la France est qu'il y a des acteurs privés, semi-privés et publics qui ont apporté à un moment donné de la valeur dans la chaîne globale.

G. Chartron fait le constat qu'il y a très peu de chercheurs présents à ce colloque et le regrette car les acteurs autour de la table ne peuvent se substituer aux besoins des chercheurs.

Elle fait remarquer que le droit des plateformes a souvent été abordé mais pas celui de la propriété intellectuelle. La qualification de la publication reste un enjeu et si on estime qu'il est nécessaire d'avoir des acteurs et regards tiers à l'égard de ce que produisent les chercheurs au niveau de la publication, il est important de composer avec eux.

Or cela n'a pas été le cas dans le cadre de la BSN et comme l'a rappelé le SNE précédemment, il y a eu peu de dialogue en France. On a assisté à un durcissement de la part des acteurs publics et des affrontements ont eu lieu.

Pour l'agenda 2015 au sein du groupe du GFII, beaucoup des acteurs du contenant réfléchissent sur les modèles économiques et sur les transformations. Il serait important d'avoir un appui des acteurs du public sur ces transformations économiques en réfléchissant avec les acteurs du contenant : les licences nationales ont été abordées. Le GFII est sensible au soutien des acteurs nationaux.

Il y a également un vrai travail à réaliser sur les compétences, sur le savoir-faire en TDM. Au GFII, les compétences existent : le président du GFII est spécialiste du TDM (Temis). Il faudrait créer des partenariats entre privé et public et déléguer les compétences afin de ne pas réinventer la roue. Il appartient au CNRS de développer les services. Il faudrait faire en sorte que les talents nationaux puissent trouver une place dans l'agenda 2015.

Intervention de l'ADBU : Ce colloque a été organisé dans l'objectif de trouver le modus operandi entre privé et public. Or l'interprofession du livre et celle de l'IST n'ont pas les mêmes intérêts, ce qui ne veut pas dire qu'il n'est pas possible de collaborer. Il y a des dossiers où il serait intéressant de collaborer avec le secteur privé, d'autres où cela n'est pas possible dont le TDM. Le CSPLA écrit dans son rapport que le secteur public se livre à du parasitisme en faisant du TDM sur des textes à contenu sous droit : c'est difficile à accepter.

Il y a de la place pour tout le monde : le droit des plateformes évoque l'idée de trouver une articulation régulée entre acteurs publics et acteurs privés. Par ailleurs, le bilan de la politique de l'interprofession du livre depuis les années 80 n'est pas très positif : les éditeurs sont en difficulté, les libraires également, les bibliothèques voient leur budget baisser, et le lecteur ne s'y retrouve pas forcément.

Les intérêts sont parfois divergents mais on peut travailler ensemble comme le font les anglo-saxons, entre autres.

Aspect Formation :

La DIST travaille dans le cadre de Plan d'actions partagées (en particulier le PAP4) et dans le cadre de la BSN9, consacré au volet formation.

Ce segment s'articule autour de 2 axes :

1. Identification de compétences nouvelles en ISN (information scientifique numérique)
2. Identification des formations diplômantes initiales et continues en France. Cette étude très large a été confiée à un cabinet de conseil qui rendra son rapport pour l'été 2015.

Ces 2 axes de la BSN serviront à construire le plan de formation du CNRS qui sera mis en place en 2015.

Le plan de formation du CNRS se basera également sur les enquêtes commentées précédemment dans ce colloque et sur les 3 réseaux IST du CNRS (RENATIS, MEDICI et RBDD) qui ont fourni un travail important de programmation des demandes de formation dans le cadre de plans pluriannuels, à la demande des instances du CNRS.

Ces formations sont celles qui se déroulent dans le cadre de collaborations inter-réseaux, voire inter-établissements.

D. Torny dresse un premier bilan.

Il faut admettre le caractère asymétrique de l'agenda 2015 qui est un agenda de construction de politiques publiques et donc d'allocation des ressources humaines, financières contractuelles, par la puissance publique. Si on considère qu'il faut introduire des acteurs privés dans la discussion, cela implique que nous leur accordions une part. La question essentielle, néanmoins, reste la construction de dispositifs publics et non pas privés.

Deux conséquences :

1. Le besoin de construire une **cohérence en termes de politique nationale**. En effet, pendant les 10 dernières années, beaucoup de dispositifs ont poussé, certains en coordination, d'autres pas. Cela implique un besoin de cohérence pour la gestion.

Les exemples de Persée et d'Istex, sont très parlants. L'agenda doit trouver une cohérence pour traiter de manière équitable l'ensemble des acteurs économiques. Notamment, les partenaires réguliers situés dans l'espace européen.

Cette cohérence n'implique pas une homogénéisation, une uniformisation de l'ensemble des dispositifs.

L'agenda pourrait définir ce qui relève du cadre local et du cadre national.

2. L'entrée dans un espace de négociations : quelles sont les priorités pour l'agenda ? Certaines institutions assurent déjà des choses. Est-ce qu'on reprend des éléments déjà présents ? Est-ce qu'on introduit des éléments nouveaux ? Est-ce qu'on arrête un certain type d'éléments car ils ne sont plus prioritaires ? Ce groupe est un espace de négociation de parties prenantes et pas seulement en opposition privé public, mais également au sein du public car il y a des histoires, des objectifs des institutions différentes.

Il s'agit de trouver, comme le CNRS le fait avec la DIST, des formes de cohérences qui évitent redondance, augmentation des coûts, politiques contradictoires et abandon de projets, en incluant autour de la table des acteurs portant des intérêts contradictoires.

Les SHS ont une très grande expérience et ouverture sur la question d'avoir des stakeholders autour de la table et d'organiser la discussion pour qu'elle soit satisfaisante pour l'ensemble des parties.

Pour Jean-François Balaudé, la construction de politiques publiques gagne évidemment à se faire dans ce cadre. Pour les licences nationales, on cherche une réflexion globale sur la capacité de l'ensemble des établissements d'enseignement supérieur et des organismes de recherche de répondre à leurs besoins. Cette réponse se fera dans un cadre national qui paraît le plus à même de garantir les intérêts. Le fait que ces contraintes soient entendues par les éditeurs est un point important. Il y a bien entendu une dissymétrie dans la discussion, car les établissements sont tenus d'assurer leur mission dans le cadre de mutations très rapides. Tout cela gagne à se faire dans le cadre d'un dialogue.

Dans le cadre de l'agenda, se posent les questions juridiques, où là encore des points de vue divergents s'expriment. Mais il est entendu que des questions nouvelles surgissent désormais, dont les enjeux des données et la manière dont on peut s'en saisir sans nuire aux uns et aux autres.

L'Agenda public fixe des objectifs de politiques publiques, mais les objets sont partagés, les débats sont ouverts et ne s'imposeront pas par la force dans une décision unilatérale. Tous les acteurs reconnaissent la valeur du monde de l'édition, et y sont attachés.

Eprist rappelle que les acteurs de l'IST sont un appui aux chercheurs et les services qui seront développés dans ce cadre, dépendront de la manière dont évolue la recherche. **Il faudrait donc intégrer à l'agenda une veille sur l'évolution des pratiques de recherche**, afin que toute action se fasse au regard du besoin réel et dans le temps.

D. Torny rappelle que l'organisation des professionnels de l'IST est telle qu'ils peuvent se couper des communautés d'utilisateurs. Ce retour vers la recherche et les chercheurs est fondamental, et il faut trouver les bons représentants sous la forme d'enquêtes, de personnes ou d'institutions.

G. Chartron rappelle que **les chercheurs souhaitent des services**. La science étant globale, les chercheurs s'organisent aujourd'hui en réseaux internationaux. D'autre part, la science publique ne peut être détachée des intérêts économiques. Les chercheurs répondent aux appels à projets ANR, ou Europe avec des industriels, les raisonnements ne pourront être uniquement ceux d'une science publique.

ADBU : on a souligné la diversité du paysage public, des disciplines, des métiers, mais il faut rappeler qu'il y a une grande diversité du paysage privé.

Il y a aussi des politiques autres, comme la politique américaine autour du financement des NIH, qui à travers la constitution de dispositifs, a modifié de manière positive ce qu'était la diffusion des connaissances dans le domaine biomédical.

Il ne faut donc pas nécessairement penser national et position globale.

J.F Balaudé conclut la table ronde en insistant sur l'apparition de nouvelles questions d'ordre juridique et rappelle qu'il ne faut pas imposer de choix. La voie ne peut être que celle du dialogue, même si c'est un dialogue tendu.

Questions :

Marc Minon : En entendant la discussion, on a l'impression qu'il y a d'une part le secteur public qui met en avant une nécessité de recherche de cohérence et d'autre part, un lieu d'affrontement et de discussions plus ou moins tendus avec les éditeurs.

Or ces journées ont montré qu'avec le numérique, les choses sont devenues plus compliquées. L'ensemble de ces questions étaient discutées au sein de BSN.

L'Agenda 2015 doit être quelque chose de différent, donc de plus ouvert, mais pas uniquement aux maisons d'édition. En effet, les universités ont une fonction de recherche mais aussi une fonction de divulgation, d'enseignement des responsabilités. Il devrait y avoir dans ces discussions d'autres acteurs numériques impliqués, tels que Thomson-Reuters, et des acteurs sociaux qui utilisent les résultats de la recherche, mais qui ne sont ni éditeurs, ni membres des organismes et universités.

Il faudra retracer l'ensemble des moyens et des budgets. Il ne s'agit pas uniquement du budget des acquisitions des bibliothèques universitaires, mais de l'ensemble des moyens disponibles

pour l'IST qui doit être mis à plat afin d'étudier leur affectation dans l'intérêt de l'ensemble des chercheurs et de la population.

Définir le rôle respectif de chacun : si on n'est pas d'accord aujourd'hui sur ce qui est efficient, nécessaire, pertinent, légitime que chacun fasse, il sera difficile de progresser.

Didier Pélaprat : L'ensemble des partenaires privés attend de la clarté de la part du service public de recherche et de l'enseignement supérieur.

Nous sommes un service public de recherche sous la tutelle de pouvoirs publics, dans nos missions il y a également, comme cela a été rappelé, la diffusion des résultats de recherche à l'ensemble de la société.

Il est nécessaire d'avoir une réflexion sur ce processus de publication et de diffusion des connaissances. Est-ce que l'on délègue ? Ou bien est-ce un service public ? Si cette action est déléguée, cela entre dans la procédure de délégation de service public.

Il faut avoir le courage de décider comment en France on envisage de remplir les missions de service public de recherche, sans quoi, les autres acteurs n'auront pas de points de repères.

Alain ABECASSIS, Chef du service de la coordination stratégique et des territoires-MENESR

Alain Abécassis remercie les organisateurs pour cette nouvelle rencontre et remercie également les représentants du secteur privé d'avoir été présents et d'avoir permis de souder la communauté.

Il rappelle deux principes qui doivent guider les préoccupations et les actions, l'implication et le rôle des chercheurs et quelques enjeux de méthodes :

1. **La diffusion de l'IST doit rester sous la maîtrise de la communauté scientifique.** C'est un principe que JF Bach rappelle dans ses commentaires au rapport de l'Académie des Sciences sur l'IST. Les chercheurs doivent s'approprier, l'évolution, le devenir, l'orientation de l'IST dans sa diffusion et dans sa production. Les chercheurs ont la responsabilité des choix qu'ils exercent dans leur politique de publication, dans l'organisation et dans l'accès à l'information. Ils sont producteurs, créateurs et diffuseurs de science.
2. **L'IST, vu son enjeu stratégique, doit être portée au plus haut niveau de l'ensemble de nos institutions (universités, organismes de recherche ou responsables ministériels).** Elle ne peut être que l'affaire des spécialistes de l'IST. Elle revêt également une **importance budgétaire** qui est loin d'être négligeable. Les arbitrages budgétaires qu'elle suscite, ne doivent pas être vécus et pensés uniquement à l'intérieur des budgets des services d'IST, ou de documentation. Ils doivent être pensés à l'échelle des budgets globaux des organismes et des structures auxquels nous appartenons. Ces dépenses de l'IST sont nécessaires, tout ce qui se fait dans ce domaine a un coût mais ce coût doit être apprécié à l'aune et à l'échelle des moyens budgétaires globaux que l'on consacre à la Recherche et à l'Enseignement Supérieur.

Les initiatives sont foisonnantes et les travaux des ateliers de ce colloque le montrent bien. Il y a une compétition extraordinaire pour plus de coopération, collaboration, plus d'ouverture et de partenariats. **Il y a aussi des besoins de lisibilité et de clarté.**

Deux éléments rendent les choses très compliquées :

1. **Les phénomènes sont très mouvants et évolutifs.** Le paysage se transforme sous nos yeux avec beaucoup d'incertitudes que ce soit de la part des acteurs publics ou privés
2. Ces sujets ne peuvent s'appréhender qu'à **l'échelle internationale.** Il faut décrypter des tendances dont on a des connaissances parcellaires et qui sont à l'œuvre à l'international et dont il est parfois difficile de dégager des lignes de forces ou des tendances.

Priorités de l'année 2015 et chantiers à conforter, engager ou amplifier

1. Donner un coup d'accélérateur important **aux politiques d'Open Access, d'archives ouvertes et d'éditions publiques**. A ce titre, la politique d'archives ouvertes engagée par HAL représente une avancée et une priorité importante d'autant plus considérée la modestie des moyens mis à contribution et la sobriété avec laquelle les équipes du CCSD mènent la politique d'archives ouvertes. Ce n'est pas à la hauteur de la priorité politique qui est d'afficher une politique d'archives ouvertes. Le CCSD a su se situer à la fois à l'échelle internationale, à l'échelle nationale et aussi en proposant une politique d'archives à l'échelle des sites. Nous avons la chance de détenir avec HAL, un atout formidable et performant ayant une capacité de pouvoir couvrir tous les degrés de granularités des politiques d'archives ouvertes. Les moyens nécessaires à cette politique gagneraient à être augmentés, en regard de l'échelle que nous affichons et des enjeux qu'elle porte pour l'ensemble de l'ESR. Il ne s'agit pas uniquement d'une politique d'Archives Ouvertes mais également de la manière dont nous pourrions organiser les services apportés aux chercheurs, aux différentes institutions, à leur visibilité dans une compétition internationale accrue. Les **politiques d'incitation ou d'encouragement sous forme de journaux, d'épi-journaux** sont des éléments essentiels qui devront être accompagnés.

Raymond Bérard a insisté sur la **mise en réseau des différentes structures** qui peuvent porter des enjeux de politique d'archives ouvertes ou de publications, d'Open Access. Cette coopération entre l'Inist, le CCSD, le Cléo, Persée, et d'autres, est un élément important. La grande force de l'IST est de savoir travailler en réseau tout en intégrant les politiques de site.

2. **Les éditeurs ont un rôle à jouer** dans les deux volets que forment les politiques d'archives. Istex présente avec Hal (archives ouvertes), la caractéristique commune de devoir concevoir une plateforme de services de qualité adressée aux chercheurs et aux institutions. Tous les investissements, toutes les imaginations qui seront mises au service du projet Istex sont autant d'éléments déterminants pour positionner la France de manière originale et à un niveau de qualité et d'ambition élevé dans l'IST avec des choix d'archives innovants. ISTEEX est une plateforme très largement ouverte à tous les chercheurs qui tient compte des spécificités disciplinaires impliquant des pratiques différentes. Elle permet d'extraire des principes communs tout en intégrant des modalités différentes de mise en œuvre.

En effet, la fonction documentaire des SHS appelle des traitements qui ne seront pas identiques à ceux d'autres domaines. ISTEEX devra en tenir compte. Il y a eu une première politique et des vagues d'achat prometteuses réalisées dans ce cadre. Elles font apparaître les limites et les besoins concurrentiels des éditeurs scientifiques français en SHS. Ceux-ci ne semblent pas avoir reçu, sur la base des sondages et des questionnaires adressés aux communautés de recherche, des réponses aux attentes et besoins qui paraissent nécessaires. Le comité de pilotage d'Istex tente aujourd'hui de prendre en compte le dynamisme des éditeurs de SHS en France. **Les réalisations d'Istex représentent la deuxième grande priorité de l'agenda 2015**

3. **Les licences nationales** permettent un accès large à la documentation et impliquent de mettre d'accord les acteurs du paysage qui sont multiples. Il y a des dispositions contractuelles à mettre au point pour avoir une licence nationale dont la définition du périmètre d'établissements concernés au-delà des établissements relevant du seul MENESR. En effet, il y a des établissements de l'ESR qui ne relèvent pas de ce Ministère et qui peuvent avoir besoin de cette ressource.

La loi du 22 juillet 2013 prévoit une co-tutelle du MENESR sur l'ensemble des établissements de l'ESR quel que soit le Ministère dont ils relèvent. La licence nationale permet d'assurer, à tous les chercheurs, l'accès le plus large aux ressources documentaires. Cela implique des difficultés d'ordre budgétaire, juridique et de périmètres, qui ne doivent pas être sous-estimées.

La politique du MENESR se veut la plus ouverte possible en matière de recherche qu'elle se réalise dans le secteur privé ou public. La responsabilité politique globale pour notre pays de devoir assurer une vitalité de recherche et de développement, montre que l'on doit être aussi attentif au dynamisme et aux progrès réalisés dans les entreprises et le secteur privé qu'à ce qui se réalise dans les établissements placés sous la tutelle du MENESR.

La notion de licence nationale et d'accès le plus ouvert possible à la recherche participe à cette philosophie et traduit la **volonté de rendre la science le plus accessible possible**.

Modalités pour rendre cet objectif de science ouverte possible et réaliste

1. **Définir de manière transparente et concertée des périmètres** de bénéficiaires des acquisitions qui sont menées sur fond public. Pour cela il faut discuter avec les éditeurs pour trouver des points d'équilibre, des compromis pour définir quel est le périmètre des ayants droits pour éviter de payer plus que ce qui est raisonnable afin que les nouvelles acquisitions n'assèchent pas des perspectives d'expansion de dynamisme et de marché pour des éditeurs nationaux et internationaux. Ces derniers ont des surfaces et des enjeux de marché qui leur permettent de savoir bien mesurer leurs limites dans les négociations que nous partageons.
2. **L'utilisation et l'exploitation à des fins de recherche et des fins pédagogiques, des contenus acquis.** Les moyens numériques permettent des éléments d'investigation et de rapprochement puissants à partir des plateformes. On peut penser que ces outils vont entraîner des progrès considérables de la science et des chercheurs. Il faudra sécuriser juridiquement les conditions d'exploitation des contenus légitimement acquis. Certains pays procèdent dans ces domaines-là par voie législative ou réglementaire, d'autres peuvent procéder par des négociations de clauses contractuelles. Quelles que soient les voies, les conditions juridiques devront être consolidées d'autant que le domaine est évolutif et les points d'équilibre trouvés à un moment donné peuvent être redéfinis en fonction des évolutions des pratiques ou des technologies. Cette **sécurité juridique est autant nécessaire pour les éditeurs que pour les acheteurs publics**. Il faut avoir la possibilité de répondre aux attentes de la communauté scientifique et aux demandes et à l'imagination des chercheurs.

3. **Les réunions d'échange comme celle d'aujourd'hui** sont importantes et offrent **une tribune au Ministère**, lui permettant de se positionner par rapport à la définition d'une politique d'IST. Parmi les outils qu'il soutient, la BSN, s'est imposée dans le paysage comme le cadre dans lequel tous les acteurs de la communauté scientifique se reconnaissent. Le MENESR a ainsi su proposer un cadre non interventionniste, sans imposer d'orientations mais en offrant des cadres permettant des co-conceptions, des co-organisations pour la définition de politiques de recherche et d'enseignement supérieur en particulier en ce qui concerne la politique documentaire. Cette infrastructure permet de concilier l'appartenance très forte au service public et à ses valeurs, et en même temps une très forte autonomie et une forte ambition des établissements, des institutions et des structures qui composent le paysage de l'ESR. La BSN est ainsi un exemple tout à fait abouti et avancé d'un point d'équilibre entre un cadre national et une responsabilité des différents acteurs de l'IST.

Priorité de l'agenda 2015 : renforcer la coopération et la coordination entre les différents outils de mutualisation qui interviennent au service de l'IST. Les réunions comme celle d'aujourd'hui participent pleinement à cet objectif, car elles aident par le dialogue et l'échange entre les différents acteurs. Elles permettent de mettre en œuvre des points d'équilibre définis par une communauté. Ces points d'équilibre bénéficient alors de l'accompagnement de l'Etat. Le MENESR répond, ainsi, aux invitations des entités telles que le CNRS (DIST ou Inist), Couperin, Abes, ou d'autres, pour échanger et travailler sur ces domaines, pour accompagner les politiques qui auront été définies et pour aider à les définir et à hiérarchiser les priorités. Celles qui doivent être portées à l'échelle nationale car définies à l'échelle internationale et européenne. **Cette complexité d'organisation, cette énergie, ce temps des palabres** est le prix à payer pour la complexité de notre paysage et de nos institutions. Par le passé, l'existence d'organismes de recherche puissants, d'universités dynamiques et puissantes a été un atout pour notre tissu. Il ne faut pas y renoncer mais il nous faut en maîtriser la gestion et la complexité pour éviter les doublons et la déperdition d'énergie alors que les moyens publics sont rares.

Conclusions générales du colloque

Renaud FABRE
DIST- CNRS

Renaud Fabre remercie Alain Abécassis pour son soutien et son engagement.

Il souhaite faire le point sur ce qui va pouvoir être fait ensemble et annonce la création d'un groupe d'expert nationaux.

Lors de la première rencontre en mars 2014, sur l'innovation et la gouvernance de l'IST, tous les acteurs concernés avaient alors eu le souci de la mise en place d'une **orchestration collective**. Cette deuxième rencontre a permis de progresser sur cette voie.

Le groupe de travail pourra travailler et se constituer comme vecteur de liaison à partir des actes de ce colloque, des comptes rendus des trois ateliers, de tout ce qui est sorti de la discussion collective.

Un certain nombre de personnes a donné son accord pour travailler ensemble au sein de ce groupe dans les semaines qui viennent afin de bâtir l'agenda 2015. Il pourra être présenté à la DGESIP et au DGRI pour ensuite être communiqué, pour information, au copil BSN. Au travers de discussions, de partage, il permettra d'avancer tous ensemble, sans omettre les difficultés, en accueillant les différences, et en faisant en sorte que nous ayons des horizons communs (ce qui est déjà le cas). Ces quelques propositions seront préparées par un groupe composé de :

Présidents : Jean-François Balaudé (Université Paris Ouest), Renaud Fabre (CNRS)

Représentants de l'Industrie de l'information : Charles Huot (GFII), Olivier Dumon (Elsevier), Jean-Marc Quilbé (EDP sciences), François Gèze (SNE et la Découverte), Marc Minon (Cairn)

Représentants des Communautés de Recherche : Raymond Bérard (CNRS), Claude Kirchner (Inria), Didier Torny (CNRS), Anne Cambon-Thomsen (CNRS), Didier Pélaprat (INSERM), + VP (CPU)

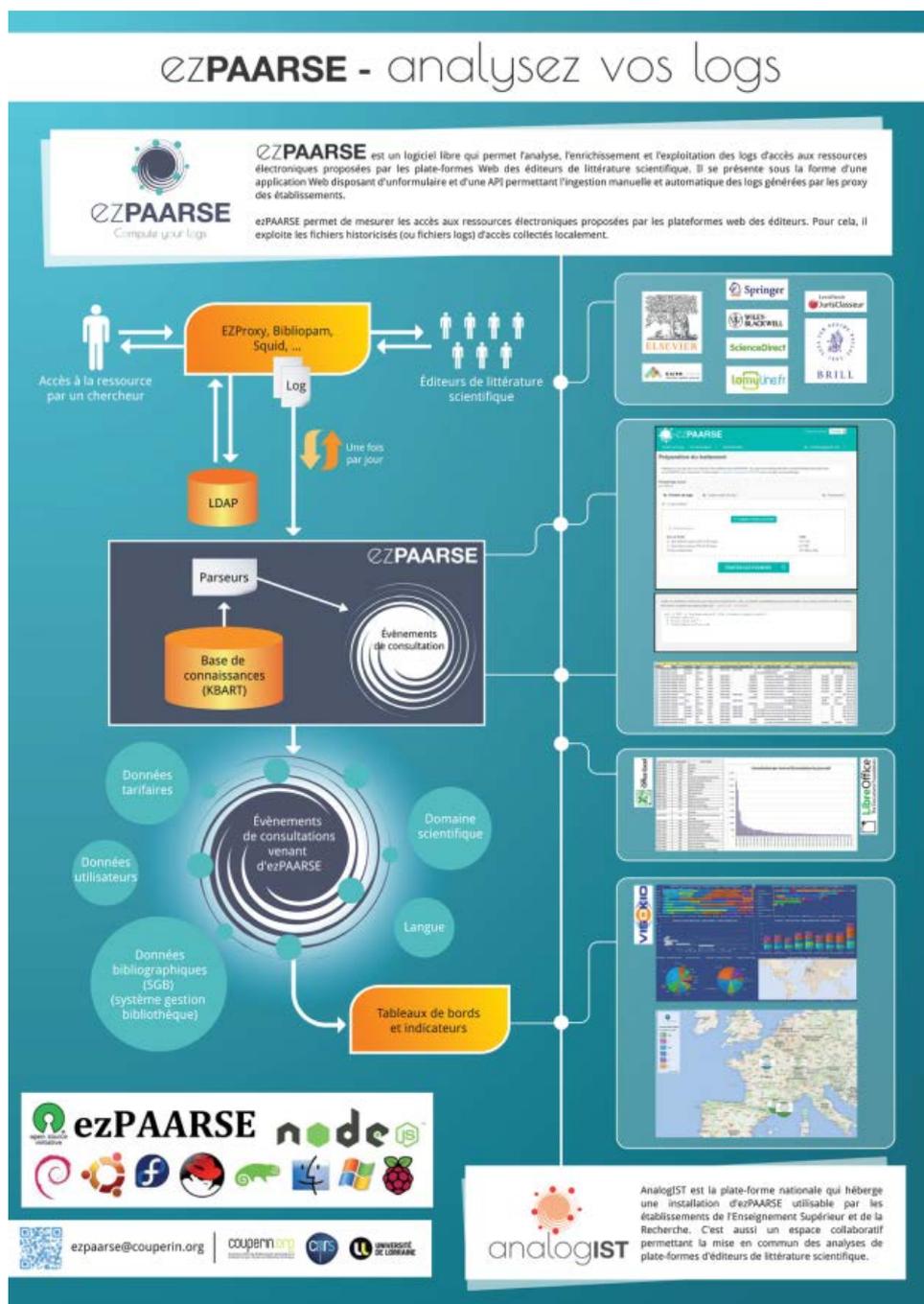
Cela augure d'un certain nombre de travaux en commun dans les mois qui viennent : il s'agira d'éviter la multiplication des rencontres pour laisser place à tous les travaux techniques dont nous avons besoin.

Le travail de ce groupe d'experts nationaux sera relayé par le site de la DIST et par la lettre CorlST.

Présentation d'ezPAARSE

ezPAARSE est un logiciel libre qui permet l'analyse, l'enrichissement et l'exploitation des logs d'accès aux ressources électroniques proposées par les plateformes Web des éditeurs de littérature scientifique. Il se présente sous la forme d'une application Web disposant d'un formulaire et d'une API permettant l'ingestion manuelle et automatique des logs générés par les proxy des établissements.

ezPAARSE permet de mesurer les accès aux ressources électroniques proposées par les plateformes web des éditeurs. Pour cela, il exploite les fichiers historicisés (ou fichiers logs) d'accès collectés localement.



Liste des participants

Nom	Prénom	Adresse mail	Organisme
ABECASSIS	Alain	alain.abecassis@recherche.gouv.fr	MENESR
ANDRE	Francis	francis.andre@cnrs-dir.fr	CNRS
AUTARD	Charlotte	charlotte.autard@cnrs-dir.fr	CNRS
BABINI	Dominique	babini@clacso.edu.ar	CLACSO
BALAUDE	Jean-François	presidence@u-paris10.fr	CPU - Université Paris X
BAPTISTE	Philippe	philippe.baptiste@cnrs-dir.fr	CNRS
BAUIN	Serge	serge.bauin@cnrs.fr	USPC et CNRS
BEAUCERF	Pierrette	pierrette.beaucerf@inserm.fr	INSERM
BELLOT	Patrice	patrice.bellot@univ-amu.fr	Aix-Marseille Université - LSIS
BENSOUSSAN	Alain	alain-bensoussan@alain-bensoussan.com	Cabinet Bensoussan
BEQUET	Gaëlle	gaelle.bequet@issn.org	Centre international ISSN
BERTHAUD	Christine	christine.berthaud@ccsd.cnrs.fr	CNRS CCSD
BLACHE	Catherine	cblache@sne.fr	Syndicat National de l'Édition
BOUCHE	Thierry	thierry.bouche@ujf-grenoble.fr	université de Grenoble 1
BOURCIER	Danièle	daniele.bourcier@cnrs.fr	COMETS
BRACHET-DUCOS	Corinne	brachet-ducos@orange.fr	CNRS
BUFFETEAU	Annie	annie.buffeteau-hejblum@ifpen.fr	IFP Energies nouvelles
CAMBON THOMSEN	Anne	anne.cambon-thomsen@univ-tlse3.fr	CNRS
CAVANAGH	Jean Frank	gerie.morris@reedelsevier.com	Reed Elsevier
CHAPOY	Elise	elise.chapoy@sciencespo.fr	FNSP
CHARTRON	Ghislaine	ghislaine.chartron@cnam.fr	CNAM
CHAUDOREILLE	Hélène	helene.chaudoreille@univ-psl.fr	Paris Sciences et Lettres
CHOURREU	Pierre	pierre.chourreu@univ-tlse3.fr	Université Toulouse 3 Paul Sabatier
COCQUEREZ	Jean-Pierre	jean-pierre.cocquerez@hds.utc.fr	CNRS-INS2I
COLCANAP	Grégory	gregory.colcanap@univ-evry.fr	COUPERIN - BU Evry Val d'Essonne
COLLEDGE	Lisa	l.colledge@elsevier.com	ELSEVIER
CONTAT	Odile	odile.contat@cnrs-dir.fr	InSHS - CNRS
CORET	Annie	annie.coret@cnrs-dir.fr	CNRS
DA FONSECA E SOUSA	Valérie	valerie.da-fonseca-e-sousa@cnrs-dir.fr	CNRS - DIST
DACOS	Marin	marin.dacos@openedition.org	CNRS - Centre pour l'édition électronique ouverte
DANDURAND	Caroline	caroline.dandurand@versailles.inra.fr	Inra
DASSA	Michèle	michele.dassa@cnrs-dir.fr	CNRS - INSHS
DAVID	Bruno	bruno.david@cnrs-dir.fr	CNRS
DAZY	M		
DE KOK	Véronique	veronique.de-kok@recherche.gouv.fr	MENESR - DISTRD
DE LAITRE	Virginie	virginie.de_laitre@agroparistech.fr	AgroParisTech
DE SAINT LEGER	Mathilde	mathilde.de-saint-leger@cnrs-dir.fr	CNRS
DE VATHAIRE	Jean-Baptiste	jb.devathaire@cairn.info	Cairn.info
DEHEE	Yannick	dehee@nouveau-monde.net	Numérique Premium
DENIAUD	Nadine	nadine.deniaud@mnhn.fr	Bibliothèque centrale du Muséum d'histoire naturelle
DIERS	Brigitte	brigitte.diers@cnrs-dir.fr	CNRS-INC
DUCHAMP	Cyril	cyril.duchamp@aef.info	AEF
DUCHANGE	Nathalie	nathalie.duchange@inserm.fr	Inserm
DUERMEIJER	Charon (Estelle)	C.Duermeijer@elsevier.com	ELSEVIER

DUMON	Olivier	O.Dumon@elsevier.com	ELSEVIER
EL KHOURI	Laurence	laurence.elkhouri@cnrs-dir.fr	CNRS
FABRE	Renaud	renaud.fabre@cnrs-dir.fr	CNRS
FABRY	Cécilia	cecilia.fabry@inist.fr	INIST CNRS
FARGIER	Nathalie	nathalie.fargier@persee.fr	Persée
FINANCE	Jean-Pierre	jean-pierre.finance@univ-lorraine.fr	CPU (Conférence des Présidents d'Université)
FLEURET	Etienne	fleuret@inist.fr	Inist-CNRS
FRANCOIS	Claire	claire.francois@inist.fr	INIST
GAUDOIN	Marie-Jeanne	marie-jeanne.gaudoin@cnrs-dir.fr	CNRS
GEZE	François	f.geze@editionsладécouverte.com	Syndicat National de l'Édition
GIRARD	Françoise	francoise.girard@polytechnique.edu	Ecole Polytechnique et CNRS
GREMILLET	Anne-Solweig	anne-solweig.gremillet@univ-paris3.fr	CNRS-Creda
GROUDIEV	Stéphanie	stephanie.groudiev@insee.fr	INSEE
GULLY	Stéphane	stephane.gully@inist.fr	INIST CNRS
HARTMANN	Silke	silke.hartmann@copernicus.org	Copernicus Publications
HERSH	Gemma	g.hersh@elsevier.com	ELSEVIER
HOLOGNE	Odile	odile.hologne@versailles.inra.fr	INRA
HUOT	Charles		GFII
JACQUEMOT- PERBAL	Marie-Christine	marie-christine.jacquemot@inist.fr	INIST-CNRS
JANIK	Joanna	joanna.janik@cnrs-dir.fr	CNRS - DIST
JANNES OBER	Emmanuelle	emmanuelle.jannes-ober@irstea.fr	EPRIST -IRSTEA
JOLY	Monique	monique.joly@insa-lyon.fr	INSA de Lyon - SCD Doc'INSA
KIRCHNER	Didier	claud.kirchner@inria.fr	Inria
KNELLER	Elisabeth	elisabeth.kneller@math.u-psud.fr	CNRS
KOERS	Hilke	h.koers@elsevier.com	ELSEVIER
LADANT	Julie	julie.ladant@recherche.gouv.fr	MENESR
LE BLANC	Annie	annie.le-blanc@cea.fr	CEA
LECOQC	Fabrice	lecoqc@inist.fr	INIST-CNRS
LEGRET	Sylvie	sylvie.legret@ifsttar.fr	IFSTTAR
LEMAITRE	Francis	lemaitre@msh-paris.fr	FMSH (ESCoM-AAR)
LITZLER	Pauline	pauline.litzler@inist.fr	INIST-CNRS
LOWINGER	Hélène	helene.lowinger@inria.fr	Inria
LUGINBUHL	Albertine	albertineluginbuhl@orange.fr	KARGER PUBLISHERS
MARCHAND	Mondane	mondane.marchand@thomsonreuters.com	THOMSON REUTERS
MARCIL	Myriam	myriam.marcil@univ-poitiers.fr	Université de Poitiers - SCD
MARIAN	Michel	michel.marian@recherche.gouv.fr	MENESR
MARTIN	Caroline	Caroline.martin@irstea.fr	Irstea
MASSON	Emilie	emilie.masson@cnrs.fr	CIL CNRS
MINIER	Arnaud	minier@numeriquepremium.com	Numerique Premium
MINON	Marc	marc.minon@cairn.info	CAIRN.info
MIN-PICAULT	Sophie	sophie.min-picault@ujf-grenoble.fr	Cellule Mathdoc
NIVARD	Jacqueline	Jacqueline.Nivard@ehess.fr	EHESS
NOQUET	Nicolas	nicolas.noquet@thomsonreuters.com	Thomson Reuters
OLLENDORFF	Christine	christine.ollendorff@ensam.eu	Ensam
OMATA-SOLAND	Izumi	izumi.omata@inserm.fr	INSERM
PELAPRAT	Didier	didier.pelaprat@inserm.fr	Inserm
PERALES	Christophe	president@adbu.fr	ADBU
PETALAS	Ilias	iliaspetalas@free.fr	CNRS
PLASZCZYNSKI	Stéphane	plaszczynski@lal.in2p3.fr	CNRS

POTELLE	Stéphane	directeur.executif@adbu.fr	ADBU
POUPON	Catherine	catherine.poupon-czysz@inist.fr	Inist
POUYLLAU	Stéphane	stephane.pouyllau@huma-num.fr	TGIR Huma-Num
PUEBLA	Iratxe	ipuebla@plos.org	PLOS
QUILBE	Jean-Marc	Jean-marc.quilbe@edpsciences.org	EDP sciences
RAMANANA-RAHARY	Suzy	suzy.ramanana-rahary@cnrs-dir.fr	CNRS/DASTR
RAYMOND-DENISE	Agnes	agnes.raymond-denise@pasteur.fr	Institut Pasteur
RENAULT	Stéphane	renault@mmsch.univ-aix.fr	LAMPA - UMR7269
RIVALLE	Guillaume	guillaume.rivaille@thomsonreuters.com	Thomson Reuters
ROTA	Anne-Catherine	a.rota@elsevier.com	ELSEVIER
ROUSSEAU-HANS	Françoise	francoise.rousseau@cea.fr	CEA
ROUSSEL	Clotilde	clotilde.rousseau@inist.fr	INIST
SAUDER	Laurent	laurent.sauder@inist.fr	INIST-CNRS
SCHEIDER	Isabelle	isabelle.scheider@univ-lyon1.fr	Université Claude Bernard Lyon 1
SIGOGNEAU	Anne	anne.sigogneau@cnrs-dir.fr	CNRS -DASTR
SORGER	Christoph	christoph.sorger@cnrs-dir.fr	CNRS-INSMI
STEFFANN	Sylvie	sylvie.steffann@inist.fr	INIST
STROPPA	Yvan	yvan.stroppa@univ-orleans.fr	LEO - UMR7322
TAKEDA	Kenneth	kenneth.takeda@cnrs-dir.fr	CNRS INSB
TALOUR	Pascale	pascale.talour@ujf-grenoble.fr	CNRS - OSUG
THIEULIN	Benoît	benoit.thieulin@cnumerique.fr	Conseil National du Numérique
THOUVENIN	Nicolas	nicolas.thouvenin@inist.fr	INIST-CNRS
TISSEAU DES ESCOTAIS	Aleth	aleth.tisseau-des-escotais@uvsq.fr	DBIST de l'UVSQ
TORNY	Didier	didier.torny@cnrs-dir.fr	CNRS
TOUMBOU	Ousseni	ousseni.toumbou@cnrs-dir.fr	CNRS
VAN TIGGELEN	Barend	barend.van-tiggelen@cnrs-dir.fr	CNRS-INP
VERDIER	Henri	henri.verdier@modernisation.gouv.fr	ETALAB
WEIL-MIKO	Christine	weilmiko@inist.fr	DIST-CNRS
WOLF	Dominique	dominique.wolf@unistra.fr	Université de Strasbourg
ZASADZINSKI	Alain	alain.zasadzinski@inist.fr	INIST-CNRS
ZIKRA	Marie-Claude	zikra@univ-paris1.fr	CNRS