



**HAL**  
open science

## Pour une meilleure visibilité de la recherche française

Antoine Blanchard, Elifsu Sabuncu

► **To cite this version:**

Antoine Blanchard, Elifsu Sabuncu. Pour une meilleure visibilité de la recherche française. [0] Deuxième labo. 2015. hal-01251541

**HAL Id: hal-01251541**

**<https://hal.science/hal-01251541>**

Submitted on 6 Jan 2016

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

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



Distributed under a Creative Commons Attribution 4.0 International License

Pour une meilleure visibilité de la  
recherche française

# Pour une meilleure visibilité de la recherche française

Livre blanc

Antoine Blanchard et Elifsu Sabuncu

Deuxième labo

Paris



Pour une meilleure visibilité de la recherche française de [Antoine Blanchard et Elifsu Sabuncu](#) est mis à disposition selon les termes de la [licence Creative Commons Attribution 4.0](#) Ceci peut être votre [site web principal](#) ou [la page d'informations vous concernant](#) sur une plate forme d'hébergement, comme [Flickr Commons](#)., except where otherwise noted

# Contenu

5 INTRODUCTION

7 RECOMMANDATIONS

Accès aux données sur la recherche

Gestion des informations tout au long du cycle de vie de la recherche

Annuaire de chercheurs et valorisation de l'expertise des laboratoires

35 COMMENTAIRES ET RÉACTIONS

39 BIBLIOGRAPHIE

40 À PROPOS DES AUTEURS

41 À PROPOS DE DEUXIÈME LABO

42 MENTIONS LÉGALES

# Introduction

## La recherche française, combien de divisions ?

La formule de Staline se demandant quelle était la force de frappe du Vatican pourrait très bien s'appliquer aujourd'hui à la recherche française. Qui comprend encore quelles thématiques scientifiques sont prioritaires ? Qui fait quoi dans le millefeuille des LabEx, EquipEx, IdEx, unités de recherche, COMUE ex-PRES, Instituts du CNRS, Alliances thématiques... ? Qui possède une vision année après année des financements accordés par l'Agence nationale de la recherche (ANR) ? Comment savoir sur quels projets travaille tel chercheur et avec quels résultats, ou qui est spécialiste de tel sujet dans telle université ? À toutes ces questions légitimes, portées par les décideurs, les chercheurs ou les citoyens, il n'y a pas de réponses évidentes et surtout pas d'endroit où les chercher.

Comme l'affirmait le recteur de l'université de Liège, Bernard Rentier, aux [Journées Couperin de janvier 2013](#), un dirigeant de la recherche doit savoir ce que ses labos produisent au même titre qu'un chef d'entreprise doit savoir quelles pièces ses employés fabriquent. De ce point de vue, les tableaux de bord de la recherche française sont bien lacunaires. La France pilote sa recherche dans le brouillard et les scientifiques en sont rendus à suivre le sens changeant du vent et à user d'entregent politique pour conduire au mieux leurs travaux. Quant aux citoyens, ils essayent de suivre... et sont frustrés de n'avoir accès qu'à la partie émergée de l'iceberg, celle qui est médiatisée, sans possibilité de savoir de quoi la recherche française est réellement faite.

Or ce n'est pas une fatalité. Pour l'université de Liège, il a suffi de [rendre obligatoire le dépôt des publications des chercheurs dans son archive ouverte](#) pour percevoir l'étendue des recherches menées, et la quantité des travaux publiés (qu'ils n'imaginaient pas). Et pour un pays comme la France ? C'est pour répondre à cette question que nous avons rédigé ce livre blanc.

## Méthode et objectifs

Afin de comprendre d'où venait le manque de visibilité et de lisibilité de la recherche française, nous avons épluché un grand nombre de documents, rapports et articles ; testé de nombreux logiciels ; et interrogé une douzaine d'acteurs du monde de la recherche. Nous avons ainsi pénétré dans les rouages de l'administration, du pilotage et de la valorisation de la recherche.

À défaut d'avoir pu tout couvrir, l'enquête nous a menés notamment sur la voie de l'*open data*, mais pas au sens où on l'entend traditionnellement en science : ce qui nous intéresse ici ce sont les données sur la recherche récoltées par les administrateurs, et non pas les données scientifiques produites par les chercheurs. Voici quel sera notre cheminement :

### Pour un accès aux données de la recherche

Dans le premier chapitre, nous arguons que des données sur l'activité et la production de la recherche existent déjà mais qu'elles sont enfouies dans les organismes. Une première étape consisterait donc à les rassembler et les publier : c'est [l'injonction de l'open data faite à l'administration française](#) et qu'est chargé de coordonner un service du Premier ministre, Etalab. Pourtant, nos contacts avec l'Agence nationale de la recherche et le Ministère chargé de la recherche ont montré que nous étions loin du compte.

### Pour une meilleure gestion des informations tout au long du cycle de vie de la recherche

Après avoir mis en lumière l'absence de données permettant de repérer les projets de recherche en cours et de suivre le chemin menant des financements aux résultats, plusieurs lecteurs nous ont fait remarquer que les chercheurs sont beaucoup sollicités "pour fournir ce genre de données, avec des formats différents et toujours avec des questions posées un peu différemment ; souvent aussi avec des informations stupides et inutiles et qui prennent beaucoup de

temps à agréger” (Gaïa). Il ne faudrait pas non plus que nos préconisations conduisent les chercheurs à passer encore plus de temps “à entrer des données dans [leurs] systèmes administratifs au détriment du temps dédié à la recherche” (Gilles Saint-Martin).

Dans ce deuxième chapitre, nous montrons qu’une bonne gestion des informations tout au long du cycle de vie de la recherche permet justement de limiter la déperdition d’information et les saisies multiples. Tout ce qui est saisi une fois ne doit jamais être resaisi une deuxième fois, mais bien plutôt réutilisé autant de fois que nécessaire. Ce que rendent possible des formats standards et des logiciels modernes, comme le montre l’exemple de la Grande-Bretagne !

### Annuaire de chercheurs et valorisation de l’expertise des laboratoires

Dans le troisième chapitre, nous clôturons notre enquête avec une question lancinante posée aux universités et laboratoires français : sur quoi travaillent vos chercheurs ? Qui est spécialiste en logique formelle dans votre institution ? Avec qui dois-je collaborer pour mon projet industriel sur les métamatériaux ? Questions qui nous ramènent en partie à l’accès aux données sur la recherche, et à la gestion des informations tout au long du cycle de vie de la recherche, vus précédemment.

## Un livre blanc pour passer à l’action

Notre analyse devrait rassurer les chercheurs, qui se plaignent souvent à la fois du manque d’information sur les orientations de la recherche et de la difficulté à repérer les bons interlocuteurs sur tel ou tel sujet, et des lourdeurs administratives qui les poussent à renseigner plusieurs fois des formulaires équivalents. La contradiction n’est qu’apparente : il existe des solutions pour éviter les pertes d’information, tout en maximisant l’usage et l’accès à cette information.

Notre analyse devrait également rassurer les administrateurs : il est possible de rendre l’administration de la recherche conviviale et directement utile, en retirant toutes les corvées (ou les tâches perçues comme telles).

Notre analyse devrait enfin rassurer les dirigeants : ce que nous décrivons n’est pas un idéal sorti de nos cerveaux mais des processus, des infrastructures déjà éprouvés en Grande-Bretagne et ailleurs... avec une analyse coûts-bénéfices qui ne laisse aucun doute quant à l’opportunité de rejoindre le mouvement !

Chercheurs, administrateurs, dirigeants de la recherche : la feuille de route est claire et la balle dans votre camp...

# Recommandations

1. Créer une association professionnelle des personnels administratifs de la recherche (comme [ARMA](#) en Grande-Bretagne)
2. Doter la Mission de l'information scientifique et technique et du réseau documentaire (MISTRD) du Ministère de l'éducation nationale, de l'enseignement supérieur et de la recherche d'une fonction d'animation par appel à projets, comme le JISC en Grande-Bretagne, qui a permis à de nombreux projets pilotes de voir le jour, et de développer les collaborations public-privé
3. Débloquer le projet CAPLAB en résolvant le conflit qui oppose le CNRS et l'AMUE, et lancer au plus vite l'appel d'offres compétitif pour sélectionner la meilleure solution CRIS pour le marché français
4. Publier la feuille de route du Ministère sur le pilotage de la recherche *evidence based*, suite au séminaire « Empirical Foundations of Science and Innovation Policy » [organisé en septembre 2013](#)
5. Confier au Conseil stratégique de la recherche une mission de réflexion sur la standardisation des remontées d'information sur la recherche, [à l'image de son homologue allemand le Wissenschaftsrat](#)
6. Mettre en place un projet pilote portant sur la remontée des informations concernant les publications et l'impact des projets de recherche financés par l'ANR
7. Amener l'ANR à libérer ses données sur les dotations qu'elle octroie et à les rendre disponible sur le [portail data.gouv.fr](#)
8. Définir un format standard de CV de chercheur ([par exemple ORCID](#) ou [SciENCv](#)) pour simplifier les réponses aux appels à projets de recherche nationaux
9. Doter le groupe de travail BSN3 des moyens suffisants pour faire aboutir le projet Conditor dans les meilleurs délais
10. Rendre obligatoire le dépôt des publications des chercheurs français dans les archives ouvertes institutionnelles et relier ces publications au futur « référentiel de la production scientifique française » du projet Conditor
11. Diffuser la norme VIVO pour rendre interopérables les annuaires de recherche en France et construire un méta-annuaire national comme [VIVOsearch](#), [DIRECT2experts](#) ou [CTSAssearch](#)
12. Intégrer le consortium CASRAI pour rejoindre le concert des nations en matière d'interopérabilité des informations de recherche

# Accès aux données sur la recherche

## L'open data selon l'ANR

Forts du mandat donné à Etalab pour favoriser l'ouverture des données des établissements publics administratifs (dont fait partie l'ANR), nous avons demandé en mars 2012 que soit publié sous forme exploitable l'ensemble des données concernant les projets financés par l'agence (montants, laboratoires, chercheurs, thématiques, durée...). Cette demande portée par Deuxième labo était co-signée par sept chercheurs, développeurs et journalistes. En voici un extrait :

Bonjour,

nous (...) demandons solennellement à Etalab l'ouverture des données de l'Agence nationale de la recherche, en particulier celles qui concernent les dotations qu'elle octroie.

Depuis la Loi d'orientation et de programmation de la recherche de 2006, l'Agence nationale de la recherche (ANR) regroupe plusieurs aides de l'État auparavant dispersées, pour offrir aux chercheurs un guichet unique permettant de financer leurs projets de recherche. Les moyens d'action de l'ANR, même s'ils sont réduits en comparaison de l'enveloppe générale de la recherche, sont des moyens non récurrents qui jouent un rôle important dans les orientations de recherche des laboratoires. On estime que l'enveloppe de l'ANR représente environ la moitié des financements incitatifs reçus par la recherche publique, ce qui correspond à ce jour à 4,8 milliards d'euros accordés à près de 9 000 projets.

Un mois et demi plus tard nous avons été invités à réitérer notre demande sur l'«espace collaboratif d'échanges» d'Etalab, [ce que nous fimes](#). En l'absence de réaction, nous avons rencontré en janvier 2013 le webmaster de l'ANR et une chargée de mission. Qui nous ont confirmé que ce n'était pas une priorité de l'ANR, au motif que l'ANR ne représenterait que 8 à 10 % du budget total de la recherche publique (mais la moitié des financements incitatifs « sur projets », comme nous l'indiquions dans notre demande) et que le paysage de la recherche française étant complexe (on l'avait remarqué, merci), l'ANR ne saurait agir en dehors d'une action groupée.

En attendant, il faut se contenter des informations que l'ANR publie sur son site web, à la fois incomplètes et dispersées. Nous avons fait l'inventaire pour vous :

- pour chaque appel à projets, pas toujours évident à trouver, vous pouvez accéder à un fichier PDF qui contient la liste des projets sélectionnés, avec pour chacun : l'acronyme du projet, son titre et le nom du chercheur qui le porte. C'est tout. Ni résumé, ni mots clés, ni affiliation ou tutelle, etc.
- vous pouvez également faire des recherches [dans la base de données des projets](#) dite « Suivi bilan », avec un succès variable : nous n'y avons pas retrouvé le [projet FADO](#) (sélectionné en 2007) ni le [projet VSN-RAP](#) (sélectionné en 2009) ; en revanche quand ça marche, vous pouvez tomber sur [des fiches projet assez complètes](#)
- si vous n'avez pas peur de vous faire passer pour un partenaire privé (puisque c'est aux partenariats public-privé qu'il est destiné), le [Moteur de la recherche](#) donne les titres des projets financés par l'ANR et le nom des labos mais sans résumé. La navigation est sommaire et l'absence de permaliens ne permet pas de partager la page d'un projet. Plus grave sans doute, la couverture est toujours imparfaite : [les tests effectués](#) par le réseau du CNRS « Valorisation en sciences humaines et sociales » montrent que sur 12 projets, aucun n'a pu être retrouvé à partir des mots contenus dans les titres. « De plus dans certains cas seul le nom du laboratoire apparaît sans même un titre de projet. »
- enfin, les [jolis cahiers thématiques](#) que l'ANR publie au compte-goutte vous donneront une information supplémentaire, à savoir la liste des productions scientifiques de chaque projet. Cependant, il faudra vous contenter parfois de paragraphes très vagues sans référence à aucune publication. Alors que la liste précise

des publications est indispensable pour qui veut pouvoir les consulter, éventuellement complétées par des observations plus générales.

Heureusement, nous avons une bonne nouvelle : ce n'est pas une fatalité, comme le montre la rectrice Denise Pumain et le président d'université Frédéric Dardel dans leur [rapport sur l'évaluation de la recherche et le passage de l'AERES au HCERES](#) (janvier 2014). Contrairement à la France où manque l'«interopérabilité des données de base qui sont souvent dispersées ou dupliquées entre plusieurs lieux de production et d'utilisation», au Brésil «chaque chercheur et enseignant chercheur dispose d'un registre qui contient l'ensemble de ses activités et de sa production...»

Partons donc étudier quelques bonnes pratiques de l'étranger, en commençant par le Royaume-Uni !

## Pendant ce temps, au Royaume-Uni et ailleurs

La première chose qu'on attend d'une base de données sur la recherche, c'est d'être à jour et utilisable. La deuxième, c'est d'offrir les informations les plus complètes possibles, avec possibilité d'entrer par projet, par laboratoire, par thématique, par porteur... et d'agréger les données à ces différents niveaux au lieu de faire se côtoyer des informations parcellaires qui ne se parlent pas. Enfin, la troisième c'est d'être réutilisable selon les principes de l'*Open Definition*, pour laisser le champ libre à de nouvelles analyses, des mashups, des visualisations de données etc.

Tout ceci, on le trouve au Royaume-Uni.

Prenons par exemple l'organisme qui finance les projets de recherche en sciences de la vie, le BBSRC : [il propose un moteur de recherche](#) couvrant les données de 3 170 dotations octroyées pour un total de 1,2 milliard de livres, avec une date de dernière mise à jour clairement indiquée. La base de données comprend aussi bien [des fiches projet](#) que [des fiches chercheur](#) (voir capture ci-après).

### Investigator / Supervisor details

|                              |   |
|------------------------------|---|
| <b>Name</b>                  | Dr Richard Morris                                       |
| <b>Current Institution</b>   | John Innes Centre                                       |
| <b>Current Department</b>    | Computational and Systems Biology                       |
| <b>Institution website</b>   | <a href="http://www.jic.ac.uk">http://www.jic.ac.uk</a> |
| <b>Previous Institutions</b> |   |

### Current Awards

| Reference                        | Title   | Award Role             | Value £   | Type              |
|----------------------------------|---|------------------------|-----------|-------------------|
| <a href="#">BB/2018627/1</a>     | <a href="#">Deciphering the Symbiotic Calcium Code: The activation of CCaMK</a> | Principal Investigator | 727,113   | Research Grant    |
| <a href="#">BBS/E/J/00000511</a> | <a href="#">Calcium signaling and flowering time modeling</a>                   | Principal Investigator | 1,007,664 | Institute Project |
| <a href="#">BBS/E/J/0000649</a>  | <a href="#">Protein structure and dynamics - information processing</a>         | Principal Investigator | 112,889   | Institute Project |
| <a href="#">BBS/E/J/0000643</a>  | <a href="#">Modeling shape during development</a>                               | Principal Investigator | 114,540   | Institute Project |
| <a href="#">BBS/E/J/000CA496</a> | <a href="#">Deciphering the Symbiotic Calcium Code: The activation of CCaMK</a> | Principal Investigator | 3,695     | Institute Project |

### Completed Awards

| Reference                          | Title  | Award Role             | Value £ | Type              |
|------------------------------------|--|------------------------|---------|-------------------|
| <a href="#">BB/F016220/1</a>       | <a href="#">Zemike moments and descriptors for 3D object processing</a>  | Principal Investigator | 58,540  | Research Grant    |
| <a href="#">BBS/E/J/00004197</a>   | <a href="#">Dynamic Integration of Functional Genomics Data</a>  | Co-Investigator        | 146,542 | Institute Project |
| <a href="#">BBS/E/J/000CA340</a>   | <a href="#">Zemike Moments and Descriptors for 3D Object Processing Submitters Reference</a>   | Principal Investigator | 1,480   | Institute Project |
| <a href="#">BBS/S/C/2005/13209</a> | <a href="#">A Bayesian framework for integrating genomic data to aid function prediction from structure (Funded via Training Grant BB/B513927/1)</a> | Principal Supervisor   |         | Studentships      |

I accept the [terms and conditions of use](#) (opens in new window)

[export PDF file](#)

[back to list](#)

[new search](#)

Exemple de fiche d'un chercheur financé par le BBSRC

**Voici une base de données à jour et facile d'utilisation, avec des informations riches et agrégées à différents niveaux.** Mais ce n'est pas tout : le BBSRC, avec 7 autres agences de financements de la recherche et de l'innovation britannique, a mis en place le [portail national « Gateway to Research »](#) qui recense l'ensemble des subventions (toutes disciplines confondues) et leurs bénéficiaires depuis 2006 ! Un véritable état des lieux de la recherche au Royaume-Uni, en version bêta pendant 1,5 an puis [lancé officiellement en décembre 2013](#).

Ce portail possède la même structure de données que le site du BBSRC, avec 10 fois plus de données : [fiches projet](#) (voir capture ci-après), [fiches chercheur](#), et prochainement des fiches secteur géographique et des fiches thématiques.

## Multiple light input signals to the gene network of the circadian clock

Lead Research Organisation: [University of Edinburgh](#)

[← Go back](#)

[Overview](#) [Organisations](#) [People](#) [Publications](#)

### Abstract

The circadian clock is a small gene network with multiple feedback loops, which generates 24-hour biological rhythms in almost all eukaryotes and in some prokaryotes. The dynamics of the clock system are complex, because expression of multiple network components is also forced by the environmental day/night cycle via light signalling pathways. My lab published the first model of the plant clock. We have validated predictions of the model by experiments, and extended the model to account for additional data. The modelling of light inputs remains very simple, as the model was based on data for very simple conditions: 12h light: 12h dark cycles and constant light or darkness. We will now manipulate the light input more quantitatively and dynamically, altering wavelength, duration and fluence rate of light exposure in several experimental protocols. We will also manipulate the photoreceptors that sense the light signal and their target genes in the clock network by mutation, monitoring the system's response using conventional and novel LUC reporter gene methods and high-resolution RNA assays. These data will constrain a new generation of refined models that are calibrated to real fluence rates acting through particular photoreceptor pathways. Technical developments in the computing platform are proposed in order to exploit the IBM BlueGene supercomputer for this work. We will analyse the resulting models to understand the features that contribute to the observed patterns of entrainment, and their similarity to the clocks of animals and fungi. Finally, we will test predictions of the new models by additional experiments, including new protocols inspired by the physical sciences and engineering. This will deepen our understanding of the plant clock as a leading example of plant systems biology.

Funded Value:

**£678 039**

Funded Period:

mars 07 - sept. 10

Funder:

BBSRC

Project Status:

Active

Grant Category:

Research Grant

Grant Reference:

BB/E015263/1

Principal Investigator:

[Andrew John Millar](#)



Exemple de fiche projet du portail « Gateway to research »

Le portail « Gateway to research » va plus loin en autorisant la réutilisation de ses données ([placées sous l'Open Government Licence](#)) via une API, et [l'organisation d'un hackathon](#) pour inviter des développeurs à s'en emparer !

Les plus attentifs auront remarqué dans la capture précédente d'une notice « Gateway to Research » l'onglet « Publications », parfois remplacé par « Impact ». Il recense les livrables scientifiques (*research outputs*) du projet de recherche, une information cruciale pour évaluer la réussite du projet et capitaliser sur ses productions. C'est également un levier puissant à l'heure où l'impact socio-économique de la recherche doit sans cesse être démontré. Encore plus en sciences humaines et sociales : ainsi, l'agence de financement britannique de la recherche en sciences sociales et économiques (ESRC) met en avant sur son site [une rubrique « Impacts and findings »](#), dont un « [Research catalogue](#) » qui comprend plus de 100 000 publications issues de projets financés (voir capture ci-après [pour un exemple](#)). À comparer avec l'ANR, où seuls les cahiers thématiques donnent les *outputs* des projets de recherche, souvent réduits à la portion congrue.

Nous reviendrons dans le prochain chapitre sur la mécanique des flux d'information qui permet au portail fédérateur « Gateway to research » de fonctionner si parfaitement. En attendant, signalons quelques initiatives similaires dans d'autres pays :

- en Slovaquie, [Science Atlas](#) publie les données de 5 716 projets de recherche, 682 905 publications, 35 272 chercheurs, et 17 190 conférences vidéo
- en Slovaquie, [SK CRIS](#) recense 18 165 chercheurs, 10 544 projets, 1 267 organisations et 7 425 publications et autres valorisations de la recherche ([voir ces explications en français](#))

## Local Public Services and Quality of Life: An Internet Tool for Citizen Consultation to 'Join-up' Performance Assessment

Grant reference: RES-189-25-0027

[View grant details](#)

My ESRC

This page:

[Bookmark](#) [Print](#)



### Journal article details

|   |   |
|---|---|
| Title                                   | Performance measures and democracy : information effects on citizens in field and laboratory experiments  |
| Description                             | There has been a massive expansion in published information about the performance of bodies delivering public services but little research about the effects on citizens. Research on information and political participation suggests that information cues allow citizens to economize on the need for full information, influencing their perceptions and attitudes and helping them hold democratic governments to account. This article uses field and laboratory experiments with random allocation to citizens of information cues about local government performance to evaluate their effects. It finds that a cue about relatively good performance raises citizens' perceived performance and satisfaction and a cue about relatively bad performance lowers perceived performance and satisfaction. Direct effects on citizens' intention to vote for the local incumbent are not evident and the cue about good performance is more influential on the perceptions and satisfaction of citizens who already supported the incumbent. However, overall, the findings suggest that public performance information systems can be credible to citizens. Future research should assess the effects of different forms of performance information in different service contexts coming from different information sources, with the sources potentially varying in credibility to citizens. |
| URI                                     | <a href="http://dx.doi.org/10.1093/jspart/mug057">http://dx.doi.org/10.1093/jspart/mug057</a>   |
| URI Description                         | Abstract  |
| Language                                | English   |
| Primary contributor                     |   |
| Author                                  | Oliver James  |
| Keywords                                |   |
|   | public services; access to public services; democracy; public opinion; performance  |
| Additional details                      |   |
| Volume number                           | 20  |
| Is this item peer reviewed (refereed)?  | Yes   |
| ISSN                                    | 1053-1858   |
| Copyright URL                           | <a href="http://www.oxfordjournals.org/oxfordopen/">http://www.oxfordjournals.org/oxfordopen/</a>   |
| Publisher                               | Oxford University Press   |
| Date of publication                     | 01 January 2010   |
| Place of publication                    | Oxford  |
| Is this item a pre-print or post-print? | Postprint   |
| Title of journal                        | Journal of public administration research and theory  |
| Files                                   |   |
|   | <a href="#">ojames2010performancemeasures.pdf (.pdf / 351kb)</a> <a href="#">Add to My ESRC</a>   |

Exemple de publication issue d'un projet financé par l'ESRC

- aux Pays-Bas, [NARCIS](#) recense 59 427 chercheurs, 852 320 publications, 2 894 organisations et 29 077 jeux de données (un projet qui [a fêté ses 10 ans en 2014](#) !)
- en Flandre (région néerlandophone de la Belgique), [le FRIS \(Flanders Research Information Space\) Research Portal](#) publie les données de 25 779 projets de recherche, 3 596 publications, 2 083 organisations et 26 647 chercheurs. Sur chaque notice, une visualisation permet d'explorer le réseau hybride composé de ces différentes entités (voir capture ci-après).

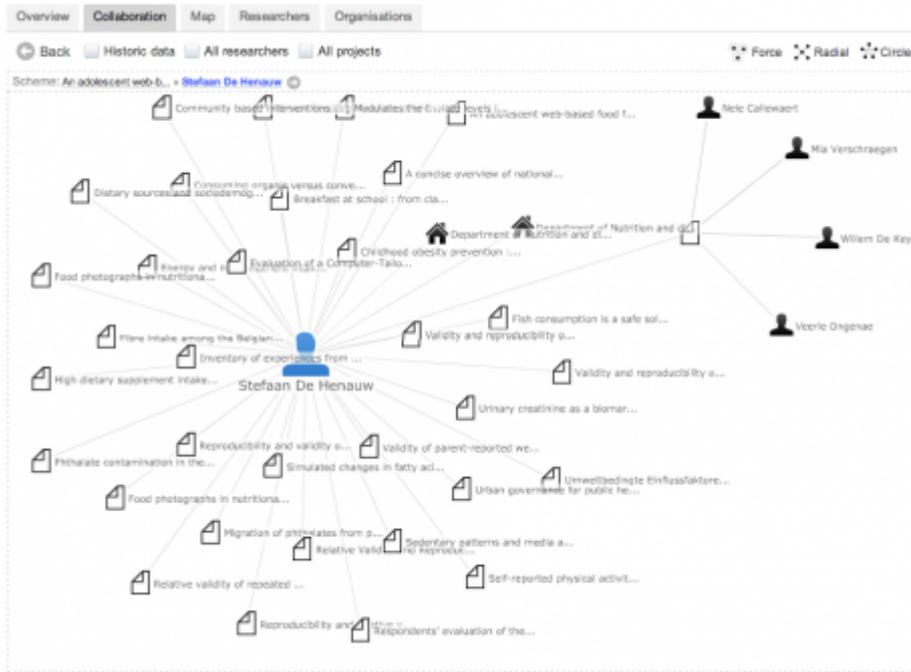
## L'open data selon le Ministère de l'enseignement supérieur et de la recherche

Gonflés à bloc par ces bonnes pratiques européennes, nous avons voulu élargir nos investigations aux données du Ministère disponible sur le portail *open data* de l'État français. Résultat en juin 2013 : rien !

C'est nous ou le @MinistereESR n'a libéré aucune donnée sur <http://t.co/a56VnQ074P> ? cc @Etalab

— Deuxième labo (@deuxieme\_lab) June 3, 2013

Passée cette surprise, [on nous encourage sur Twitter](#) à creuser un peu, jusqu'à trouver [certaines données sur les effectifs de la recherche publique](#), provenant... du Ministère de l'éducation nationale, à un moment où ces deux ministères



Réseau du chercheur Stefaan De Henauw tracé par le FRIS Research Portal

étaient séparés (voir capture ci-après) : problème de production des données ? Problème d'attribution (plus probable) ? Voilà qui ne fait pas sérieux !



Sur data.gouv.fr, les données sur les effectifs de la recherche publique proviennent du Ministère de l'éducation nationale

Interrogé, le ministère nous a répondu sur Twitter qu'il allait publier plus de statistiques ; effectivement, **en juillet 2013 était publiée sur data.gouv.fr** la nouvelle édition de l'Atlas régional des effectifs étudiants. Puis rebondissement en avril 2014 avec le lancement de la plate-forme d'exposition de données ouvertes intégrée au site du ministère (<http://data.enseignementsup-recherche.gouv.fr/>) et synchronisée avec data.gouv.fr. D'après le **rapport sur les politiques nationales de recherche et de formations supérieures** annexé au projet de loi de finances pour 2015, le ministère

a initié, à partir de septembre 2014, une démarche interne de caractérisation de son patrimoine de données sur l'ESR. Sous l'angle des données produites et disponibles et non sous celui des systèmes les produisant, il s'agit de caractériser les champs couverts, de disposer d'une vision précise de la qualité de l'information disponible, des usages qui en sont faits, des contraintes régissant leur accès ou leur diffusion. C'est sur la base des résultats de cette démarche que les directions générales définiront la programmation des nouveaux jeux de données qui seront proposés sous licence libre. Par ailleurs, d'ici à la fin 2014, le MENESR proposera à la communauté des réutilisateurs un dispositif (événement, dispositif web) facilitant l'interaction avec les producteurs des jeux de données.

**Dans cette attente, on ne peut que regretter que les jeux de données libérés soient le plus souvent des états des lieux chiffrés ou des indicateurs synthétiques, au détriment des données brutes, des référentiels etc.**

En particulier, nous sommes très intéressés par le [répertoire national des structures de recherche](#), qui est aujourd'hui en phase de pilote opérationnel après plusieurs années de discussion et de travail. Il a pour vocation de recenser l'ensemble des structures de la recherche publique déclarées par les universités, les grandes écoles ou les organismes de recherche. Il ne constitue pas un système d'information sur la recherche ; il est un référentiel qui permet de mettre en correspondance les informations sur une même structure résidant dans les différents systèmes d'information du ministère ou des établissements. Parmi les utilisateurs naturels du répertoire, le ministère citait en 2008 l'ANR, pour « tirer les bilans par établissement (universités, organismes) des candidatures et des lauréats », en rattachant chaque structure de recherche à sa ou ses tutelle(s). Ce répertoire doit aussi servir à identifier les structures pour entretenir *in fine* « une banque de données sur les compétences et les ressources des laboratoires ».

Cependant, la nécessité de conférer une certaine autorité à cette liste alourdit ses mises à jour et leur validation. Certains acteurs que nous avons interrogé préfèrent la « légèreté » d'autres répertoires « de fait », sur lesquels nous revenons au prochain chapitre. Il n'empêche que le ministère aurait vocation à rendre ce référentiel disponible au titre de l'*open data*, pour ouvrir la porte à d'autres exploitations possibles.

## Vous avez dit interopérabilité ?

Disposer de jeux de données pays par pays serait un premier grand pas. Mais autant imaginer tout de suite comment relier ces données, en partant du plus petit dénominateur commun et en échangeant nos expériences nationales. C'est l'objectif du [consortium d'origine canadienne CASRAI](#), qui réunit les financeurs de la recherche et les organismes de recherche de différents pays pour créer des standards permettant l'interopérabilité des informations de recherche. Par des ontologies, des taxonomies, etc. **ils souhaitent faciliter l'échange de données et leur réutilisation entre les équipes de recherche, les institutions de recherche et les organismes de financement tout au long du cycle de vie de la recherche.**

CASRAI a demandé à Deuxième labo d'organiser une présence française lors de la [journée d'étude CASRAI Reconnect](#) organisée le 16 mai 2014 à Rome — aux côtés de participants suédois, britanniques... déjà bien actifs. Et malgré un [appel relayé dans les médias](#), aucun établissement d'enseignement supérieur et de recherche ne s'est manifesté. À la date de février 2015, la France ne semble toujours pas représentée dans CASRAI.

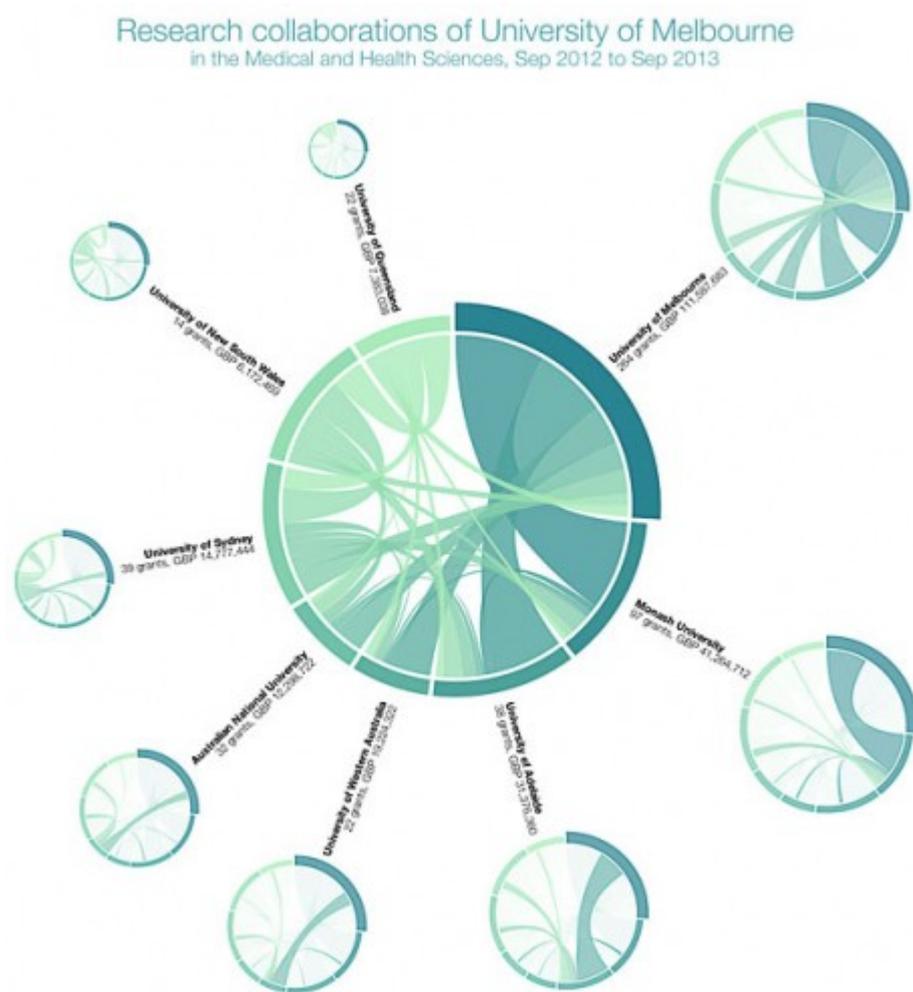
Un des objectifs de l'interopérabilité et de l'échange de données entre institutions est de permettre une meilleure compréhension du paysage de la recherche à l'échelle mondiale. Les politiques scientifiques nationales se mènent encore trop au doigt mouillé, et on ne sait encore évaluer parfaitement leur impact, ni les comparer aux politiques des voisins. Seule l'utilisation généralisée et intelligente des données sur la recherche permettra l'avènement d'une politique de la recherche fondée sur des preuves (*evidence based*).

## Vers une politique de la recherche « *evidence based* »

Quand on conduit un projet et plus généralement une stratégie, il est recommandé de l'accompagner d'indicateurs de performance permettant de mesurer si les objectifs fixés ont été atteints. Le monde de la recherche n'échappe pas aux KPI (*key performance indicators*), et ceux-ci ne sont pas toujours un gros mot. Ainsi, les organismes qui financent la recherche donnent de plus en plus d'instructions pour un libre accès aux publications : reste à en évaluer l'efficacité. Ceci nécessite d'avoir de bons systèmes d'information reliant les publications aux subventions... comme le [système centralisé FundRef](#) mis en place l'an dernier par l'organisation internationale [CrossRef](#) (celle qui assigne les identifiants DOI).

De leur côté, les administrateurs de la recherche ont des problèmes similaires : afin de mieux accompagner les équipes de recherche et d'optimiser la recherche de financements, **ils aimeraient notamment** pouvoir identifier les labos qui remportent le plus de dotations et comparer leur performance avec celles des voisins (voir vidéo ci-après). Le [projet institutionnel G4HE](#) vise à répondre à leurs besoins, en offrant un accès par lot aux données du « Gateway to Research » et des outils d'aide à l'analyse.

C'est aussi l'objectif du logiciel **Dimensions** de la société Symplectic, **lancé en septembre 2013**. Ce logiciel permet de visualiser et naviguer (voir capture ci-après) dans les données de financement de la recherche de trois pays — l'Australie (ARC et NHMRC), les États-Unis (NIH) et le Royaume-Uni (RCUK) — en attendant de futures données en provenance d'Europe, du gouvernement fédéral américain et des organismes de bienfaisance britanniques. Alors que le projet FundRef a déjà **recensé 4 000 organismes de financement de la recherche** dans le monde, l'exhaustivité des données du logiciel Dimensions semble hors de portée (au moins à ce rythme) !



Visualisation des collaborations scientifiques de l'université de Melbourne dans le logiciel Dimensions

Enfin, on peut aller plus loin avec le projet **STAR METRICS** de l'économiste américaine **Julia Lane**, qui vise à donner des fondations empiriques aux politiques de recherche et d'innovation. Les études dont on dispose actuellement dans ce

domaine sont en effet lacunaires, pas assez systématiques et pas assez fines — en particulier par manque de données permettant de retracer pas à pas les liens entre subventions de recherche, chercheurs qui les reçoivent, et production de ces chercheurs. En réunissant autour de la même table des financeurs de la recherche, des gouvernements, des opérateurs de recherche, des organisations internationales comme CrossRef ou CASRAI et des économistes, il devient possible de :

- comparer les approches et confronter les retours d'expérience
- partager le code et les méthodes
- discuter des problèmes de données et de métadonnées
- développer des approches comparatives à l'échelle internationale.

Un séminaire à ce sujet s'est tenu en septembre 2013 au Ministère de l'enseignement supérieur et de la recherche. Il réunissait quelques Français (représentant l'ANR, l'Observatoire des sciences et techniques du HCERES, le ministère et deux laboratoires d'économie de l'ENSAE et de l'université de Strasbourg) et de nombreux spécialistes étrangers dont Julia Lane et un représentant de CASRAI.

Sans tomber dans l'excès inverse de la mesure toute puissante, il nous semble que cette approche installe de bonnes pratiques de transparence et de collaboration. Puisse-t-elle sensibiliser nos institutions, toujours sensibles aux *benchmarks* et autres tableaux de bord, aux bienfaits de l'ouverture et du partage des données !

# Gestion des informations tout au long du cycle de vie de la recherche

## La recherche, un enfer administratif

C'est désormais un lieu commun de la recherche en France, et la première recommandation du [rapport de l'Académie des sciences sur les structures de la recherche publique française](#) en septembre 2012 : il faut simplifier (et les académiciens ajoutent même « arrêter de compliquer ») les procédures administratives. Explications (nous soulignons) :

La facilité de diffusion par voie électronique de questionnaires construits de manière peu rationnelle par des personnes très éloignées des laboratoires et n'ayant pas une connaissance réelle de la vie des laboratoires amène les chercheurs à passer un temps de plus en plus grand à remplir de trop nombreux formulaires qui nourrissent des « cimetières à informations » dont la taille semble seulement limitée par celle des serveurs qui hébergent ces formulaires une fois remplis. **Les chercheurs ne sont pas au service des différentes structures administratives, mais au contraire, ces structures doivent contribuer à faciliter la vie des créateurs et des innovateurs.** Faire simple au lieu de faire compliqué doit être le principe de base à respecter lors de la mise en place de toutes les modifications qui devront être apportées aux structures de la recherche publique en France. Par ailleurs, la mise en place de ces simplifications doit se faire en s'appuyant sur des avis de scientifiques compétents et des indicateurs simples, efficaces et validés.

La préoccupation des académiciens est également financière : « Dans une période de contraintes budgétaires, nous tenons à souligner que nos propositions de simplification et d'arrêt de la complication des procédures administratives sont une excellente manière de redéployer, à volume constant, des moyens financiers vers les laboratoires et les équipes de recherche ». C'est une motivation qui s'ajoute au triste constat de tant de temps gâché pour si peu de résultat ! À la clé, même pas un « [Gateway to Research](#) » à la française, comme nous le regrettons au chapitre précédent...

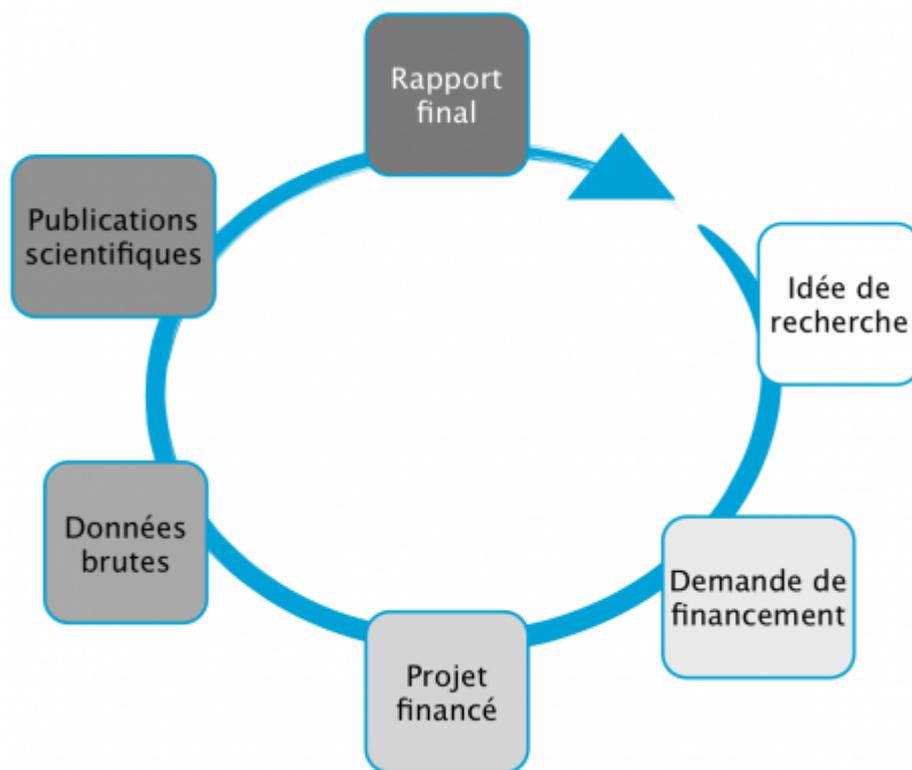
## Le cycle de vie de la recherche

Pour comprendre ce qui pourrait être amélioré dans l'administration de la recherche, il faut partir du cycle de vie (théorique) de la recherche et commencer à le décomposer en logiciels, systèmes d'information et bases de données... Un travail titanesque ! À défaut de tout couvrir, nous saurons exactement d'où vient ce sentiment de gâchis même quand des efforts sont faits pour fluidifier la circulation des informations.

Le cycle de vie de la recherche, c'est cette **circulation des connaissances d'un document à l'autre**, passant du stade de questions de recherche dans une demande de financement à celui de résultats publiés dans un article scientifique, faisant naître de nouvelles hypothèses à tester dans de nouveaux projets, générant des données venant nourrir d'autres projets, etc. (voir schéma ci-après).

Ce cycle devrait être fluide, et pourrait l'être si (par exemple) chaque information saisie était réutilisée plusieurs fois dans plusieurs contextes : « *input once, output many* ». Or en pratique, chaque étape est gérée dans un système d'information spécifique : les publications sont gérées dans des outils de gestion bibliographique et stockées dans des archives ouvertes, les données de la recherche alimentent des bases de données spécialisées ou prennent la poussière sur un disque dur, chaque agence de financement a sa propre plateforme de soumission des projets de recherche, les sites web de laboratoire ne se mettent pas à jour automatiquement, etc. Ceci suppose de **développer l'interopérabilité des systèmes avec des normes et des formats standards**.

Un gros chantier, nous ne le nions pas. Mais pas nouveau : l'interopérabilité s'impose aux établissements de recherche depuis le [décret du 2 mars 2007](#), dans ses trois dimensions :



Cycle de vie de la recherche

- interopérabilité organisationnelle qui concerne les échanges automatisés d'informations entre organisations d'un même secteur ;
- interopérabilité sémantique qui définit un langage commun à base de normes et de référentiels permettant à différentes applications d'interpréter de façon homogène les données transmises pour les réutiliser sans erreur ou perte d'information ;
- interopérabilité technique qui couvre la mise en relation des systèmes par des interfaces ouvertes, l'interconnexion des services...

Voici la théorie. Pour vous donner une idée de la situation réelle, attardons nous sur le travail multilatéral entrepris par la Bibliothèque scientifique numérique (BSN), une plateforme de coopération pour le développement et l'usage de la documentation scientifique numérique, lancée dans le cadre du [plan Besson pour le développement de l'économie numérique](#), et de la [Stratégie nationale de recherche et d'innovation](#) (SNRI). Au sein du groupe de travail n° 3 de la BSN (BSN3), intitulé « Dispositifs de signalement », est conduit le [projet Conditor](#) visant à « recenser l'ensemble de la production scientifique de la communauté Enseignement Supérieur et Recherche ». Vaste programme ! Faisant le constat d'un éparpillement indéniable entre « différents référentiels de structures, dictionnaires d'affiliations, auteurs, personnel, revues, colloques, thématiques... [qui] existent au niveau national, institutionnel, local : le besoin de référentiels communs (ou liés entre eux) est exprimé par tous ». Mais si le besoin d'interopérabilité est évident, les moyens d'y parvenir ne le sont pas et Conditor démarre sous la forme d'une expérimentation réunissant l'Agence bibliographique de l'enseignement supérieur (Abes), les directions du CNRS chargées de l'appui à la structuration territoriale de la recherche (CNRS-DASTR), de l'information scientifique et technique (CNRS-DIST) et des SHS (CNRS-INSHS), l'Institut national de l'information scientifique et technique (Inist), le Centre pour la communication scientifique directe (CCSD), les instituts de recherche agronomique (Inra) et pour le développement (IRD), Inria, le Ministère de l'enseignement supérieur et de la recherche, l'université Paris Dauphine, et l'université de Bordeaux. Tout ça pour traiter uniquement de l'aval du cycle de vie de la recherche ! Parce qu'un schéma vaut mieux qu'un long discours, la figure suivante illustre à quoi Conditor doit aboutir.

Autre exemple : les chercheurs qui déposent leurs publications dans l'archive ouverte peuvent ensuite exporter leur production dans un format directement compatible avec les rapports d'activité du CNRS ([CRAC](#) en sciences formelles et expérimentales, [RIBAC](#) en sciences humaines et sociales). L'[application IdRef](#) développée par l'Abes se base sur les métadonnées saisies dans HAL pour [créer un référentiel d'auteurs, enrichi et contrôlé par des documentalistes, et s'interfaçant avec les annuaires de chercheurs existant dans les établissements](#). Mais — car il y a un mais — les compléments de saisie à chaque acte de gestion ne sont jamais exploités : ainsi, HAL nourrit les rapports CRAC ou RIBAC mais toutes les autres informations que les chercheurs ou les administrateurs y auront mis ne reviendront jamais dans HAL.

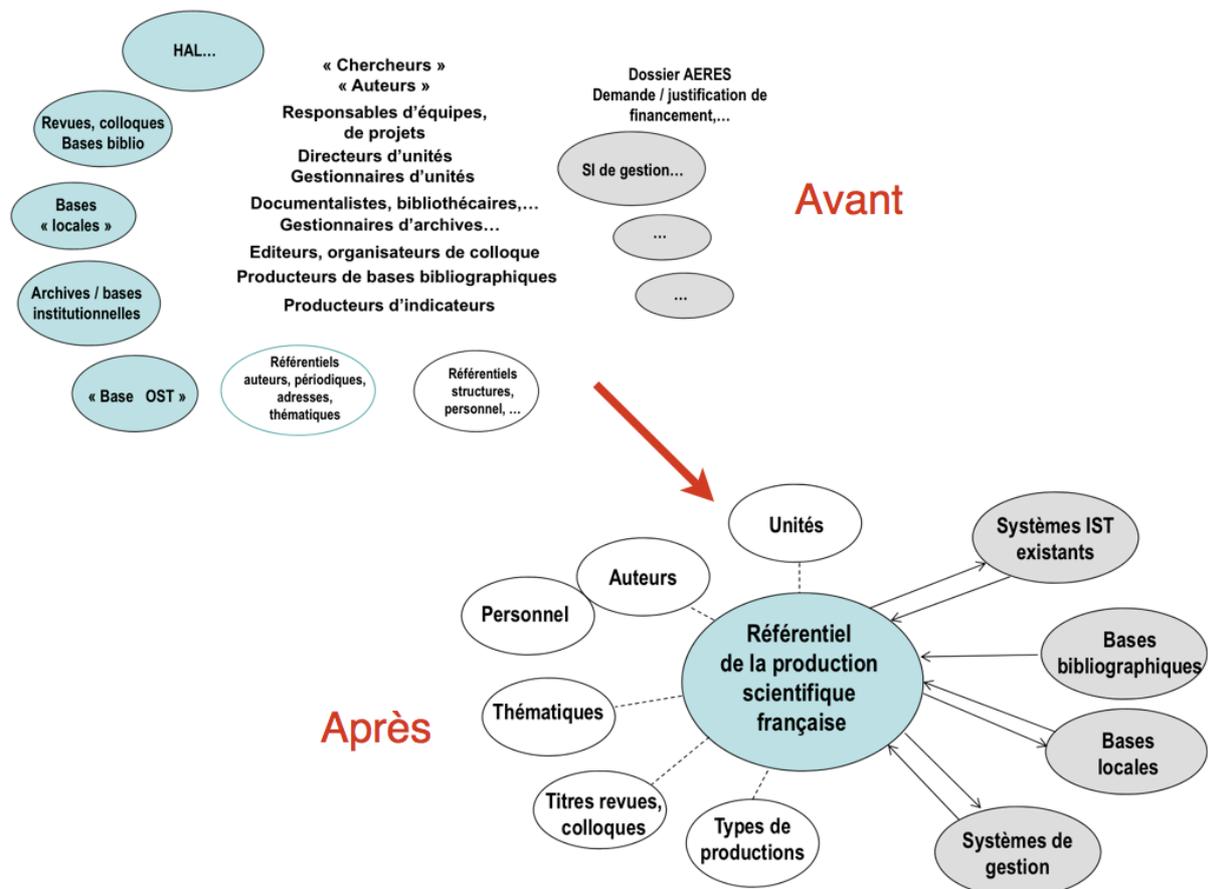


Schéma du projet Conditor, vers un référentiel de la production scientifique française. Source : Annie Coret et Raymond Berard, « Dispositifs de signalement », Journée Renatis, Cachan, 4 juillet 2012

D'où des saisies multiples, et **un sentiment de gâchis même quand des efforts sont faits pour faciliter l'échange d'informations.**

Nous voilà donc éclairés sur la situation française, et la lourde tâche consistant à fluidifier la circulation des informations. En Grande-Bretagne, **ceux qui ont abordé le problème ont identifié** « CERIF as the best option for improving interoperability and exchange of research information ». Qu'est-ce que CERIF ? C'est un des secrets derrière le portail « **Gateway to Research** », que nous allons dévoiler maintenant.

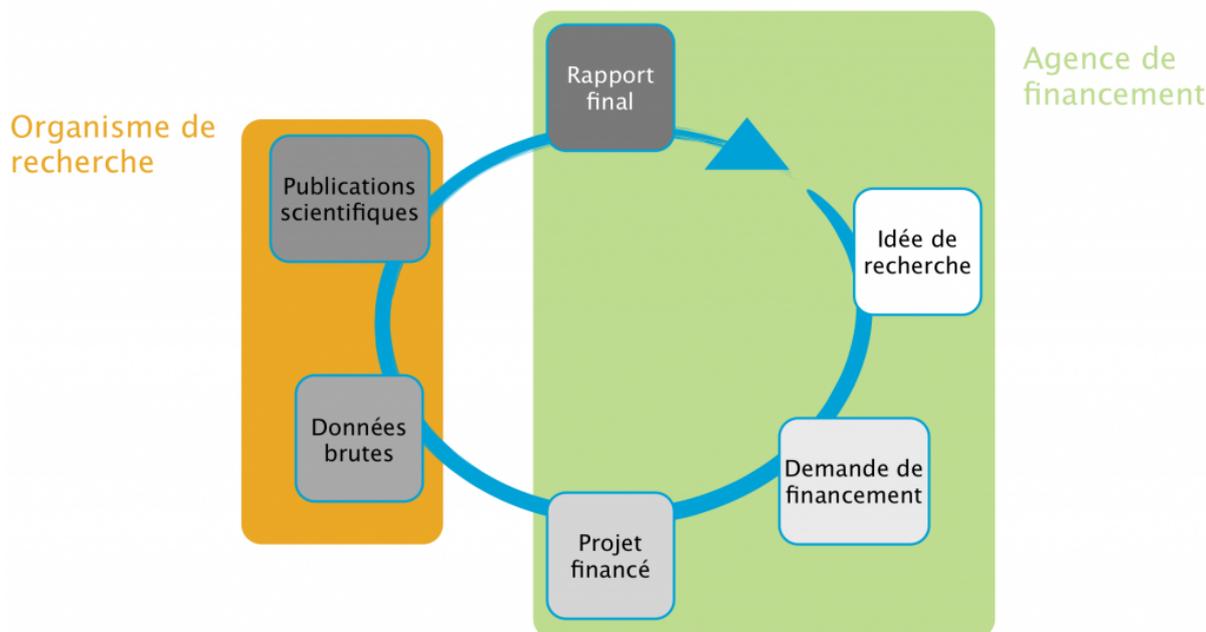
## Petit détour par l'international

Le Royaume-Uni possède avec « **Gateway to Research** » un portail fabuleux d'accès aux projets de recherche, aux chercheurs et aux laboratoires. Qu'ont-ils fait de plus que la France ? Répondre à cette question va nécessiter de pénétrer dans la mécanique des formats et standards d'échanges de données entre les laboratoires de recherche et leurs financeurs.

Reprenons notre cycle de vie de la recherche : on s'aperçoit que la gauche du cycle concerne les recherches menées au laboratoire, tandis que la partie droite concerne les agences de financement (voir schéma ci-après). Or c'est précisément sur l'optimisation des relations entre la partie gauche et la partie droite, lorsqu'organismes de recherche et agences de financement s'échangent des informations cruciales, que s'appuie le portail « **Gateway to Research** ».

Au commencement, cinq des huit agences de financement de la recherche britanniques (RCUK) ont mis en place en novembre 2011 le **système d'information ROS** (« Research Outcomes System ») pour mutualiser la saisie des informations concernant la production des recherches financées, selon neuf catégories :

- publications
- collaborations / partenariats
- co-financements obtenus
- développement des compétences du personnel
- dissémination / communication



Cycle de vie de la recherche, des agences de financement aux organismes de recherche

- propriété intellectuelle et valorisation
- distinctions et récompenses
- impact
- autres produits de la recherche

Ainsi, au lieu de maintenir cinq systèmes distincts et imparfaits, ces agences de financement ont préféré mutualiser leurs efforts pour installer un système unique, fiable et dessinant un paysage plus complet de la recherche britannique... qui alimente directement « Gateway to Research » !

En pratique, ROS peut être renseigné de trois manières différentes au moment de rendre compte de l'avancement d'une recherche (voir schéma ci-après) :

- les chercheurs peuvent saisir manuellement leurs données via les formulaires en ligne de ROS, ce qui peut vite s'avérer chronophage
- les administrateurs ont aussi la possibilité de soumettre des données par lot sous forme d'un fichier Excel, qu'ils auront exporté depuis leur système d'information interne — ce qui suppose que les champs soient bien alignés entre la sortie de ce dernier et l'entrée de ROS
- une troisième solution plus automatisée a été mise en place : l'utilisation d'un format XML interopérable, suivant le [schéma de données CERIF \(Common European Research Information Format\)](#).

Le fameux format CERIF apparaît enfin ! Les raisons d'utiliser CERIF sont nombreuses, comme nous allons le voir, et les responsables de ROS en donnent plusieurs justifications. [Nous pouvons par exemple citer celle-ci](#) (nous soulignons) :

We hope that utilising the CERIF- XML schema (...) to deliver tight data specification and set up automated data export and import routines will help save public money as **the administrative costs of delivering the information will be reduced**. Having clearer definitions will also **reduce the potential for confusion and anxiety within the academic community** and for administrators interacting with the ROS system.

CERIF représente donc un gain financier et humain évident. Au-delà même de son utilisation dans le système ROS, un [projet britannique](#) vise à déterminer dans quelle mesure CERIF pourrait servir de base à une infrastructure nationale pour le *reporting* de la recherche. Attention : « infrastructure nationale » ne veut pas dire système centralisé unique (qui peut en effrayer certains). L'existence des systèmes ROS ou autre [ResearchFish](#) n'est pas menacée, [l'objectif étant de créer des passerelles entre les différents outils et plateformes](#) — bref, de l'interopérabilité !

Or **CERIF se prête merveilleusement bien à cette ambition d'interopérabilité des informations sur la recherche, à laquelle « Gateway to Research » doit tant**. Pour deux raisons principales : parce qu'il est fait pour ça et parce qu'il est historiquement européen donc international. Nous allons détailler ces raisons dans la section suivante, assez tech-

Organisme de recherche

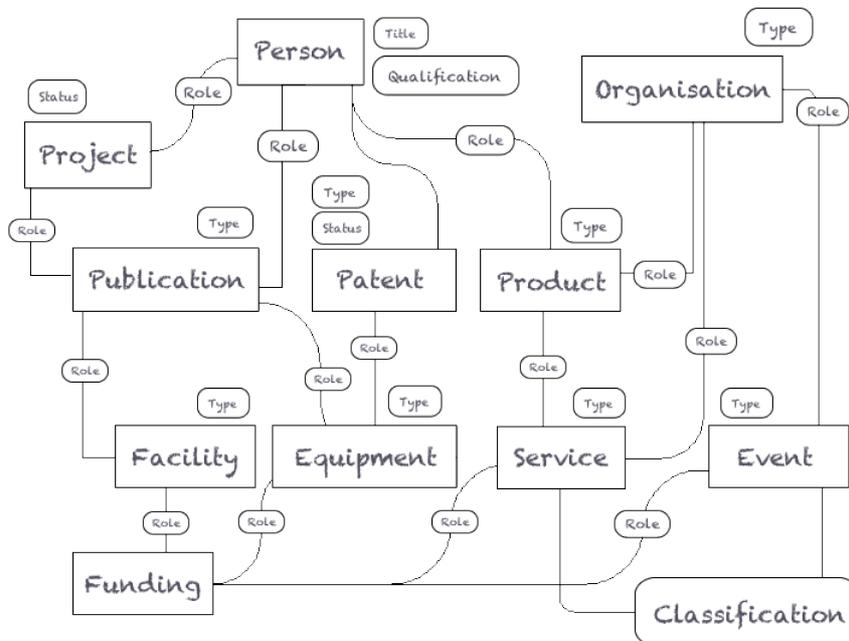


Trois modes de reporting de la recherche britannique auprès du système ROS, qui nourrit le portail « Gateway to Research »

nique, que les plus pressés peuvent sauter. Nous montrerons ensuite que CERIF fonctionne déjà dans de nombreuses autres situations concrètes.

CERIF, un cadre de référence pour l'identification des activités de la recherche

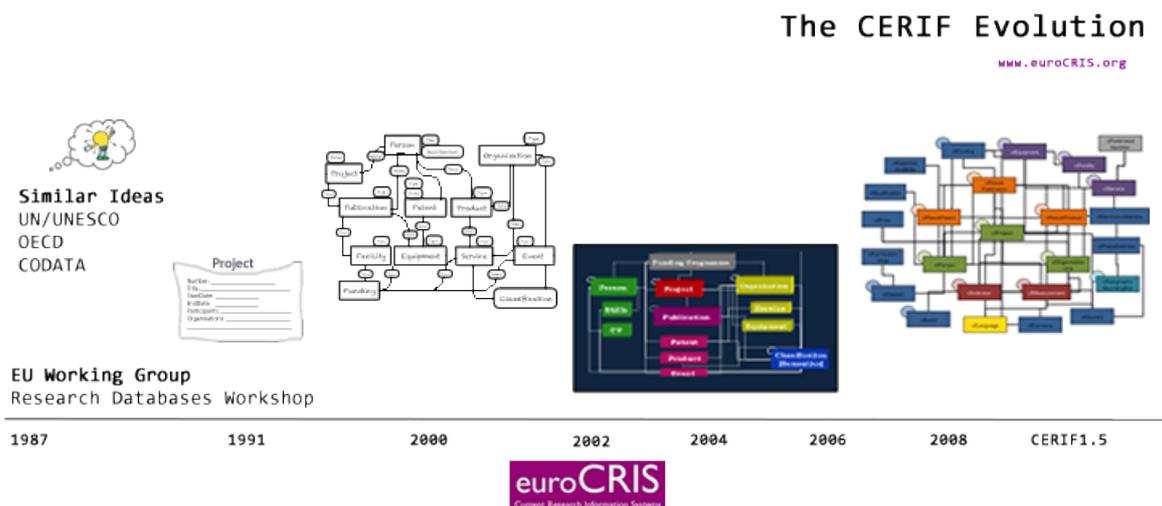
CERIF, ou « Common European Research Information Format », est un schéma conceptuel de données décrivant le domaine de la recherche sous la forme d'un modèle entité-association. À partir de ce modèle peuvent être générés des scripts SQL pour plusieurs langages de bases de données (Oracle, mySQL, DB2...). Les entités présentes dans le modèle concernent les personnes, les organisations, les projets, les publications, les brevets, les données, les équipements, les infrastructures, les financements, les indicateurs... Voici par exemple quelques unes des entités présentes dans le modèle et leurs relations :



Représentation de quelques entités et relations du modèle de données CERIF 2000. Source : euroCRIS

La structure du modèle CERIF a directement inspiré le format d'échange CERIF-XML, qui est une adaptation au langage XML d'un modèle entité-association. En ce qui concerne le web sémantique, il n'existe aucune ontologie formelle CERIF en OWL ou RDF. Des efforts ont été faits pour exprimer CERIF dans VIVO et vont être poursuivis. Néanmoins, les auteurs de CERIF insistent sur le fait que la sémantique est déjà présente dans CERIF, s'appuyant sur certains vocabulaires contrôlés, et que les ontologies du Linked Data ne feraient que la réexposer en RDF.

Par ailleurs, comme l'indique le « E » de CERIF, ce modèle est historiquement européen. Aujourd'hui fort du statut de **recommandation aux États membres de l'UE pour l'harmonisation**, CERIF a fait ses premiers pas en 1987 avec un séminaire et un groupe de travail lancés par l'Union européenne. En 1991, la première version de CERIF ne couvrait que les projets, remplacée en 2000 par un modèle de données complet avec métadonnées et modèle d'échange. En 2002, le développement de CERIF a été confié par la Commission européenne à l'association **euroCRIS** basée aux Pays-Bas. Depuis lors, les évolutions de CERIF ont permis d'introduire une couche sémantique en 2008 et le format d'échange CERIF-XML en 2012.



Frise chronologique des évolutions de CERIF. Source : euroCRIS

Notons enfin qu'euroCRIS se bat comme un beau diable pour faire reconnaître le format CERIF, en étant impliqué dans 3 projets européens : **EuroRIS-Net+**, **OpenAIREPlus** et **ENGAGE**. Malheureusement, **le manque de financement européen pour CERIF est vécu douloureusement par euroCRIS**, qui voit les États-Unis soutenir largement **leur solution VIVO** (dont on reparlera dans le troisième et dernier chapitre de notre enquête).

Mais s'il y a un pays qui a embrassé largement CERIF, c'est bien la Grande-Bretagne, avec **plus de 2 millions de livres dépensés sur le terrain** pour tester des scénarios d'usage, encourager l'adoption du format CERIF et développer ses fonctionnalités. Lors des réunions d'euroCRIS, **les autres pays montrent régulièrement une certaine jalousie** envers cette implication très forte de la Grande-Bretagne dans CERIF.

Résultat : CERIF **est devenu le standard à adopter par l'ensemble de la communauté scientifique britannique** pour « fluidifier la circulation des données entre parties prenantes, améliorer la qualité des données et réduire les coûts en permettant des procédures d'échange de données plus efficaces ». Une fois cette recommandation importante émise, **les travaux ont continué pour tester des scénarios d'usage (transfert de données à l'occasion de la mutation d'un chercheur, fonctions d'import/export entre les bases de données) et s'accorder sur un schéma CERIF-XML standard** qui satisfasse au contexte britannique tout en restant simple d'emploi. La figure ci-après illustre par exemple les bénéfices attendus du projet « CERIF in Action », par famille d'acteurs.

CERIF est donc un schéma de données standard, important pour rationaliser la circulation des informations de recherche. Mais **un schéma de données n'est rien s'il n'est pas utilisé au moment où les données sont produites. D'où l'importance des logiciels CRIS basés sur CERIF et utilisés au laboratoire et dans les organismes de recherche** pour saisir les informations liées aux nouvelles publications, aux nouveaux financements et projets, aux nouveaux personnels...

### Des CRIS pour gérer les informations au laboratoire

À ce stade le schéma de données CERIF est encore abstrait, un peu comme si on essayait de vendre la révolution bureautique en vantant les mérites du format de fichier DOC. Rentrons donc dans le cœur du sujet en s'intéressant aux CRIS, qui sont aux informations de recherche ce que les logiciels de traitement de texte sont à la bureautique. **Les CRIS forment la famille des Current Research Information Systems, ou systèmes d'information sur la recherche en cours** : ce sont des logiciels qui permettent de gérer les informations de recherche au laboratoire (personnels, projets, financements, publications...).



Bénéfices attendus du projet "CERIF in Action", par famille d'acteurs. Source : Scott Brander, « Cerif in Action: Synthesise, standardise and productionise CERIF for Higher Education Institutions », CIA Workshop, Londres, 19 octobre 2012

L'association euroCRIS gère un **répertoire des logiciels CRIS déployés dans le monde**, qui fait apparaître une forte variabilité : certains CRIS sont basés sur le modèle de données CERIF, d'autres ont leur propre modèle de données mais peuvent exporter au format CERIF-XML afin d'être compris par d'autres CRIS compatibles avec CERIF ; certains CRIS sont des logiciels commerciaux, d'autres sont des logiciels développés en interne à une université ; certains sont spécifiques à un pays, d'autres sont utilisés à travers le monde...

S'il fallait n'en retenir que quelques uns, ce seraient les **trois logiciels commerciaux les plus utilisés** :

- **Pure**, développé par la société Atira et **racheté par Elsevier** en août 2012, est un logiciel complet qui a poussé très loin dans ses dernières versions les fonctionnalités d'analyse qualitative des projets et de métrique de la recherche ;
- **Converis**, développé par la société Avedas et **racheté par Thomson Reuters** en décembre 2013, est une solution modulaire qui couvre la gestion des demandes de financement, la gestion des publications et/ou la gestion des doctorants (voir vidéo ci-après) ;
- **Elements**, développé par la société Symplectic, est centré sur la gestion des publications par les chercheurs eux mêmes, rendue très intuitive et automatisée : le logiciel envoie des alertes à chaque nouvelle publication détectée dans des bases de données comme Web of Science.

<http://www.converis5.com/res2/video/converis.mp4>

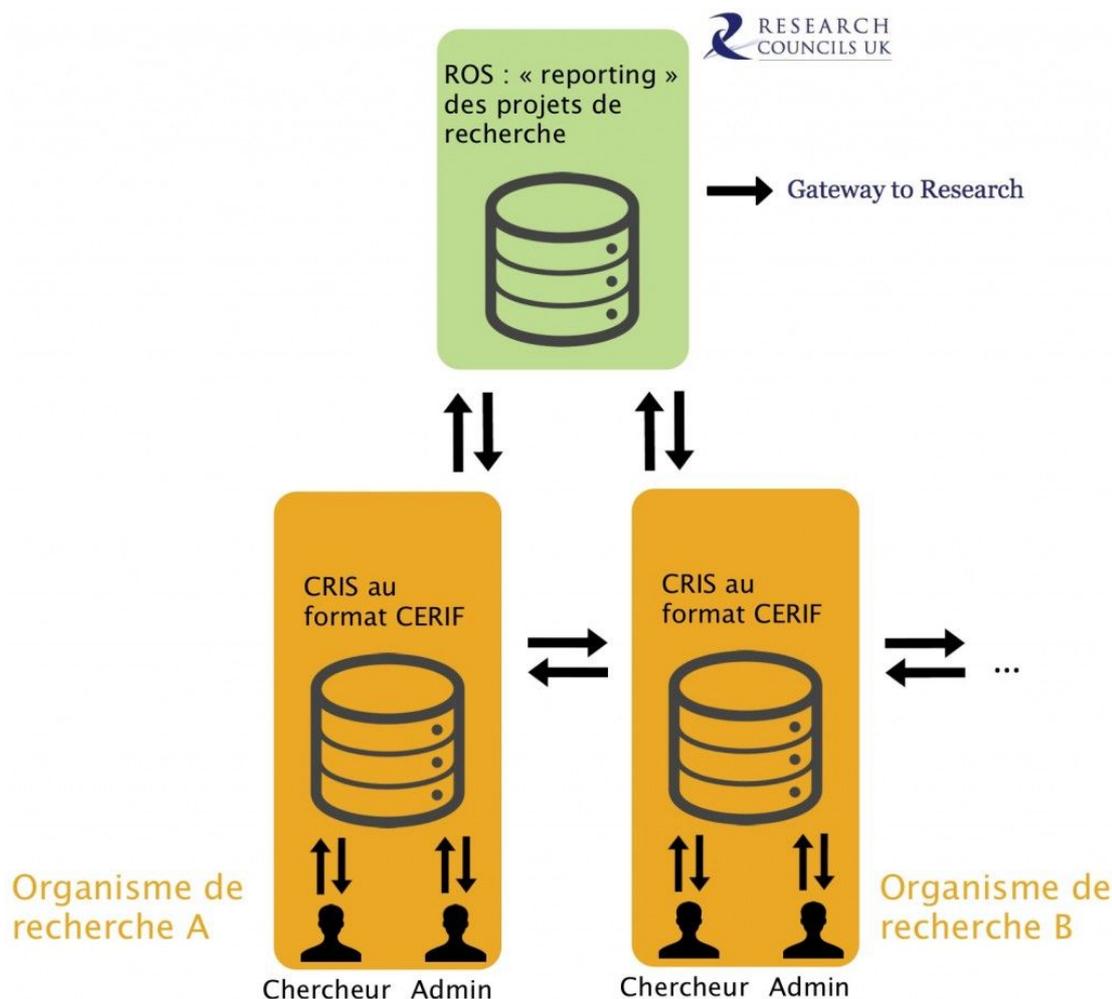
**Ces logiciels font leur trou** : on dénombre **entre 150 et 200 installations de CRIS en Europe**, et **la Grande-Bretagne est passée** de 0 à 51 installations entre mai 2009 et mars 2012, soit un taux d'adoption de 30% sur l'ensemble de ses établissements d'enseignement supérieur et de recherche. Certains pays ont fait le choix de solutions nationales, comme le **programme CRISStin en Norvège** et **METIS aux Pays-Bas**. Enfin, en janvier 2013, dans **son rapport sur la standardisation des remontées d'information sur la recherche**, le **Wissenschaftsrat** allemand (équivalent de notre **Conseil stratégique de la recherche**) a recommandé fortement que les nouveaux CRIS utilisés dans le pays **soient compatibles CERIF**.

Cette tendance n'est pas étonnante selon **une étude britannique évaluant l'économie de ce secteur**. Elle a chiffré que le déploiement de logiciels CRIS compatibles CERIF ferait économiser 177 000 livres (environ 215 000 euros) par an et par établissement, soit 25 à 30 % des dépenses annuelles liées au dépôt et à la gestion des bourses de recherche. Un système d'échange de données CERIF entre les établissements de recherche et les financeurs ferait économiser, lui, 94,5 millions de livres (environ 115 millions d'euros) par an au secteur ! Quant aux coûts liés au déploiement d'un logiciel CRIS, les retours d'expérience obtenus par cette étude convergent vers une dépense annuelle de 12 à 25 000 € par établissement.

Les logiciels CRIS compatibles avec le standard CERIF sont donc les **bricks d'un système d'information moderne, interopérable à l'échelle d'un pays et bientôt entre les pays**. Et cerise sur le gâteau, ce sont des **logiciels ergonomiques et conviviaux** qu'on aimerait avoir dans son laboratoire ou son université !

## Ce que nous apprend le cas britannique

En conclusion, on voit bien que **le portail britannique « Gateway to Research » n'est pas une réussite miraculeuse mais le résultat d'une mutualisation d'une part** (les RCUK se rassemblant pour créer le système de reporting ROS), **et d'un travail d'interopérabilité d'autre part** : les données du ROS utilisent le format CERIF, qui sert de base aux logiciels CRIS utilisés dans les organismes de recherche. Qui plus est, ces logiciels CRIS utilisés dans différents établissements leur permettent d'échanger des données, par exemple à l'occasion de la mutation d'un chercheur ou du changement de tutelle d'un laboratoire (voir schéma ci-après). Une feuille de route toute tracée pour la France... qui semble malheureusement encore loin !



Le schéma de données CERIF, un standard pour la gestion et l'échange des informations de recherche

## En France, des envies mais peu d'avancées

Pour cette dernière partie, et maintenant que nous avons compris ce qu'il était possible de faire avec les technologies disponibles en 2014, nous allons décortiquer le cas de la France. Quels CRIS utilise-t-elle ? A-t-elle dans les tuyaux des projets utilisant le potentiel de CERIF ?

Commençons par constater que la France ne peut pas feindre d'ignorer cette dynamique internationale, avec **5 Français membres de l'association euroCRIS**. Mais que pèsent 5 représentants face aux **44 membres britanniques** ? Alors que les études économiques sont si édifiantes, la France continue à utiliser des logiciels vieillissants ou en décalage avec cette vague de fond : Cocktail, Pléiade et GRAAL.

**La suite Cocktail** se définit comme un progiciel de gestion intégrée (PGI) libre pour l'enseignement supérieur et la recherche. L'un de ses modules, **SANGRIA**, est consacré à la gestion administrative et financière de la recherche à travers la gestion des structures (personnels, dotations financières, évaluations...) et des projets (réponses aux appels d'offres, partenariats...). Cocktail, né à l'IUT de la Rochelle en 1993, est **intrinsèquement lié à la communauté universitaire française**. À partir de 1997, il a commencé à être mutualisé entre les établissements intéressés, pour conduire en 2005 au consortium Cocktail chargé de définir les règles de mutualisation pour le développement des briques logicielles qui équipent aujourd'hui plus de 70 universités (sur 83) pour la gestion de leurs personnels, leurs salles, leur comptabilité... Mais **ce n'est qu'en 2011 que Cocktail a développé avec SANGRIA une solution spécifique à la recherche**. À titre d'exemple, les travaux menés dans le cadre de SANGRIA ont permis

de développer un glossaire des termes (« unité », « fédération », « tutelle », « dotation »...) utilisés pour en limiter les différentes interprétations possibles et cadrer l'utilisation de l'application.

**Pleiade est un autre PGI** mais commercial, édité par la société **QUASAR Conseil**, qui intègre le module Sirius XXL consacré à la gestion de la recherche. Celui-ci offre par exemple la possibilité d'importer une liste de publications depuis l'archive ouverte HAL pour les affecter à la production des chercheurs selon une codification conforme à celle de l'AERES. Il permet également de gérer, par contrat pluriannuel de recherche, toutes les données relatives au laboratoire, ses équipes, personnels et doctorants, ainsi que ses activités (publications, contrats de recherche, dossiers de propriété intellectuelle, prospection...). Il exporte ainsi certaines informations et tableaux exigés par l'AERES ou le ministère.

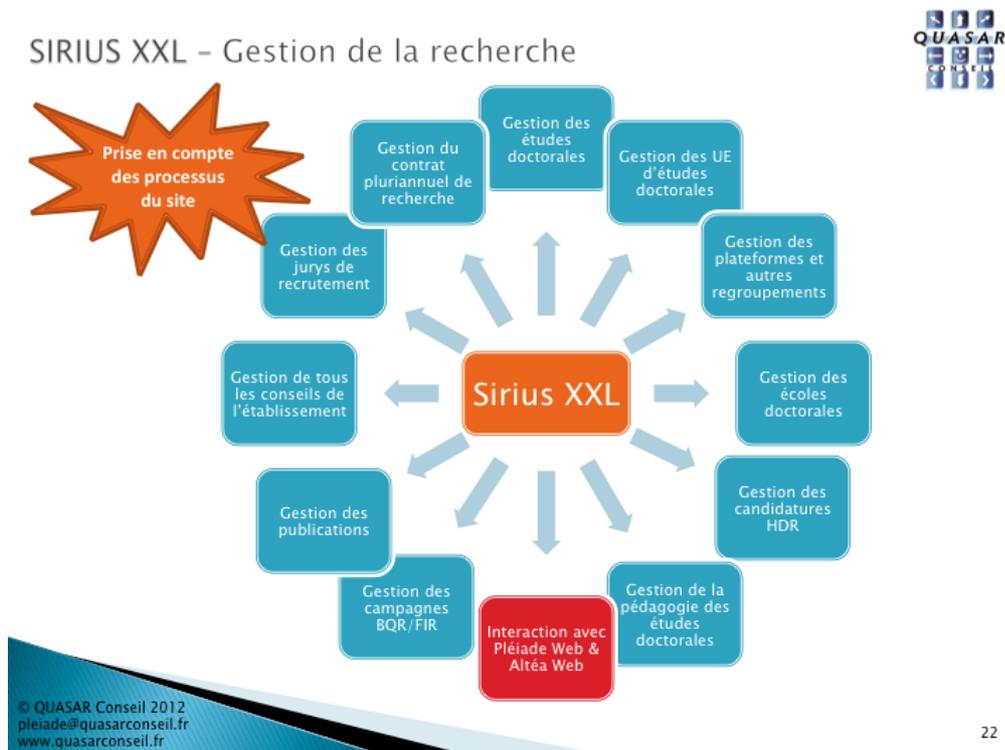
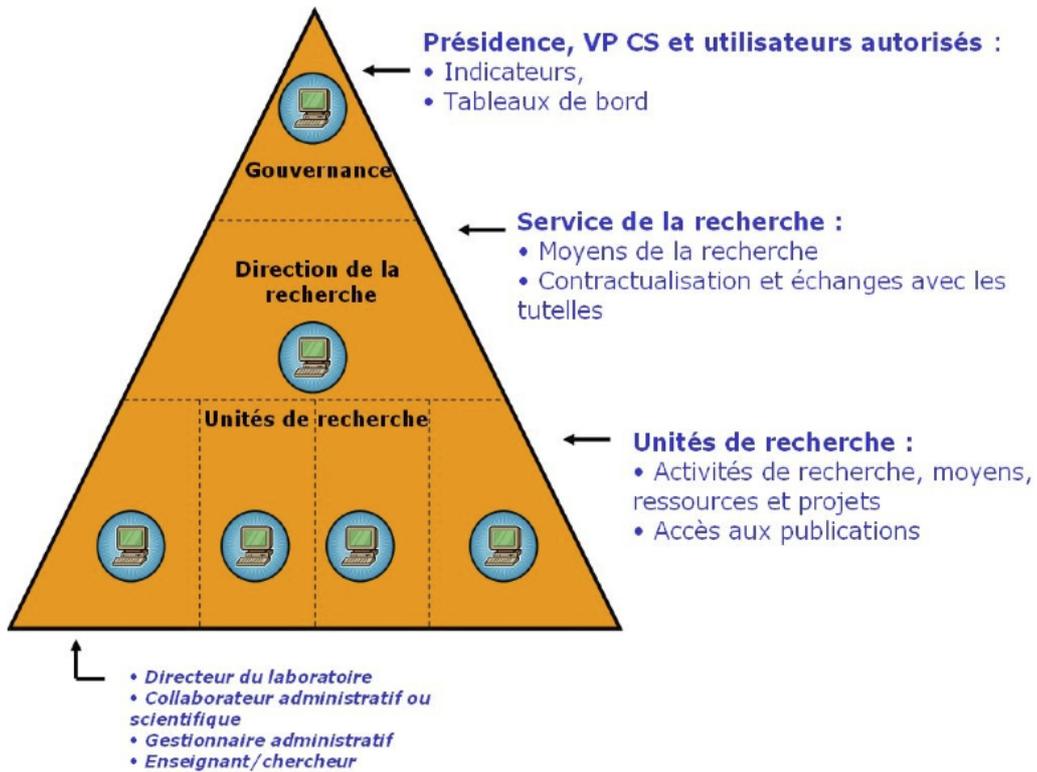


Schéma de principe de Sirius XXL. Source : « Gérer la recherche et sa valorisation avec Pleiade », QUASAR Conseil, 2012

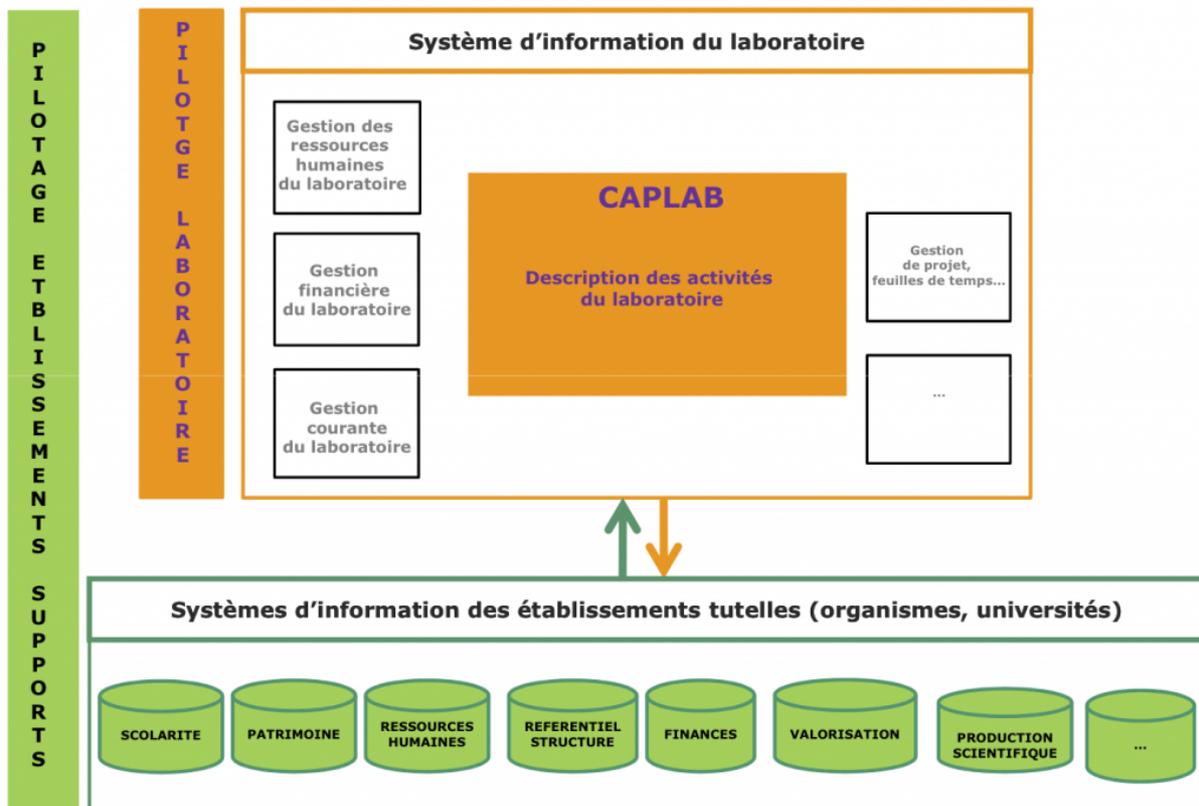
Enfin, **GRAAL (Gestion des données de la Recherche, Application des Activités Laboratoires)** est un logiciel lancé en 2000 par le Centre interuniversitaire de calcul de Grenoble puis élargi en 2004 aux quatre universités de l'académie de Grenoble, à l'université Blaise Pascal de Clermont-Ferrand et à l'université Louis Pasteur de Strasbourg. Son pilotage a alors été confié à un Groupement d'intérêt scientifique (GIS) rassemblant les partenaires publics intéressés. Fin 2009, GRAAL était déployé dans une cinquantaine d'établissements. Comme les logiciels précédents, **GRAAL est une application qui vise à présenter de manière cohérente les unités de recherche au sein de l'université**, incluant les personnels et leurs activités scientifiques (publications, brevets, manifestations, équipements...), ainsi que le suivi des moyens financiers (dotations du contrat quadriennal et sur appels d'offres) et des activités internationales (conventions de partenariat, congrès...). Comme SANGRIA, il permet d'historiciser les données et donc de suivre et gérer les transformations des unités de recherche, des moyens financiers et humains, et des indicateurs. Il permet à trois niveaux hiérarchiques dans l'établissement de saisir ou accéder à certaines données : l'unité de recherche, la direction de la recherche et la présidence (voir schéma ci-après).

Deux types d'interopérabilité existent dans GRAAL : avec SIREDO (utilisé par le Ministère dans le cadre de la gestion des contrats quadriennaux) et avec l'archive ouverte HAL. Avec HAL, l'interopérabilité existe dans les deux sens, c'est-à-dire que les métadonnées des publications peuvent être saisies dans GRAAL et attachées au fichier de l'article, l'ensemble étant ensuite versé automatiquement dans GRAAL. À l'inverse, les notices présentes dans HAL peuvent être moissonnées par GRAAL afin d'alimenter le reporting.

Mais patatras : alors que ce logiciel était bien parti pour s'imposer dans le paysage, soutenu dans son comité de pilotage par le ministère, la Conférence des présidents d'universités (CPU) et l'Agence de mutualisation des universités et des établissements (AMUE), cette dernière a annoncé en 2010 qu'elle retirait son soutien à GRAAL. Pour Michelle Gillet de la société concurrente **QUASAR Conseil**, GRAAL n'était plus à la hauteur des nouveaux enjeux de la gestion de la recherche, illustrés par le niveau d'exigence des documents à fournir à l'AERES pour la vague 2012-2015. Mais l'AMUE avait un plan B : s'allier avec le CNRS et la CPU pour **développer une solution remplaçant GRAAL, à savoir CAPLAB** (Cartographie, Activités et Pilotage du LABoratoire). Inutile de refaire le topo, résumé par le schéma ci-après :



Les trois niveaux d'utilisateurs de GRAAL. Source : AMUE, « Formation fonctionnelle de GRAAL », 2009



Principe de fonctionnement de CAPLAB, s'insérant dans le système d'information du laboratoire, en lien avec les systèmes d'information des tutelles. Source : AMUE, « Projet CAPLAB pour le description des activités du laboratoire : présentation », Club utilisateur GRAAL, 11 février 2011

Et alors qu'**aucun des logiciels précédents ne respectait le format de données standard CERIF**, l'AMUE donne des signes de bonne volonté **en adhérant fin 2012 à l'association EuroCRIS** ! Mais il y a un os : **pour l'instant CAPLAB n'est qu'une coquille vide, un projet sans logiciel**. Pourtant, **le calendrier annoncé en 2012 était sans ambiguïté** (voir capture ci-après) : après rédaction du cahier des charges, un appel d'offres aurait dû être lancé en 2012 pour concevoir et réaliser l'outil en 2013.

Nous sommes en 2015 et l'appel d'offres n'a toujours pas été lancé ! Renseignements pris en février 2013 auprès de l'AMUE, un premier *benchmark* avec tous les outils du marché a permis de restreindre le choix à 3 solutions mais seul un appel d'offres compétitif peut permettre maintenant de sélectionner l'un de ces trois prestataires.



## CAPLAB

### Un planning prévisionnel serré :

- **Etudes comparatives d'outils « recherche » similaires, retours d'expérience,...: 2012**
- **Description des processus, rédaction du cahier des charges et procédure de marché : 2012**
- **Conception et réalisation de l'outil : 2013**
- **Déploiement sur quelques sites pilote : 1<sup>er</sup> semestre 2014**
- **Déploiement sur les autres sites : à partir du 2<sup>ème</sup> semestre 2014**

Calendrier du projet CAPLAB, tel qu'annoncé en 2012. Source : Pierre Doucelance, « Point sur le partenariat avec les établissements d'enseignement supérieur et de recherche », CNRS, 9 mai 2012

Alors pourquoi tarde-t-il autant à venir ? L'explication de ce retard a été fournie par l'Inspection générale de l'administration de l'éducation nationale et de la recherche (IGAENR) dans [son rapport de décembre 2013 consacré à l'AMUE](#). On y apprend que le développement de CAPLAB a été confié à l'AMUE dans le cadre d'un accord signé avec le CNRS en décembre 2011, avant qu'apparaissent des « désaccords sur le projet de cahier des charges et sur les modalités de la consultation des éditeurs de logiciels ». [Dans son rapport annuel](#), l'AMUE explique de son côté que « la constitution d'un appel d'offres a été finalisée en septembre 2013 » mais « le comité de pilotage stratégique recherche n'ayant pu se réunir, la décision de s'engager sur la publication de ce marché a dû être reportée ».

Les inspecteurs de l'IGAENR recommandent au Ministère de l'enseignement supérieur et de la recherche de « réaffirmer la volonté de voir les projets conjoints GESTLAB et CAPLAB aboutir rapidement », et on ne peut qu'être d'accord. Mais cela suffira-t-il ? Ces atermoiements ne peuvent plus continuer, il faut désormais avancer vite !

Car **en attendant les transformations qu'ont connues des pays comme la Grande-Bretagne ou les Pays-Bas, la France en est réduite à bricoler**. Prenons par exemple les nombreuses procédures d'évaluation des unités de recherche : au CNRS, l'Institut des SHS a [lancé l'outil RIBAC en se fixant comme objectif](#) de « réduire les redondances de saisie : par exemple, l'import des références des publications entrées dans l'archive ouverte HAL-SHS ». Mais les auteurs reconnaissent qu'il reste « à progresser encore pour automatiser davantage la saisie » et la rendre plus fiable, puisque [deux personnes à plein temps sont employées](#) à vérifier certaines informations manquantes, effectuer des tâches de dédoublonnage etc. Ils aimeraient également « parvenir à une sorte de document unique, qui pourrait servir à plusieurs services, organismes, ou objectifs collectifs ou individuels » : faute de document unique, RIBAC propose actuellement un export des données en formats Word et Excel, charge aux administrateurs de les bidouiller pour les faire entrer dans les cases des autres administrations. On ne s'étonnera donc pas que le [rapport de Denise Pumain et Frédéric Dardel sur l'évaluation de la recherche et le passage de l'AERES au HCERES](#) (janvier 2014) constate que « le dossier demandé aux unités pourrait progressivement être allégé, en étant pré-rempli, à partir des bases de données du ministère ou des établissements, de nombreuses informations factuelles qu'il suffirait ensuite aux directeurs d'unités de vérifier. » Rappelons que **dans d'autres pays, c'est le format CERIF qui sert de monnaie d'échange entre les établissements, leurs tutelles, leurs évaluateurs et leurs financeurs**. Mais de CERIF, il est nul question dans ce rapport...

De ce portrait que nous avons souhaité fidèle de la situation, chacun tirera les conclusions qu'il veut. De notre côté, le choix est fait : nous demandons plus d'effort dans l'interopérabilité des systèmes d'information de la recherche française, le déploiement

de logiciels CRIS conformes aux standards internationaux, et l'accès aux données sur la recherche aujourd'hui éparpillées entre ces multiples silos... pour les exploiter à leur juste valeur !

## Bonus : de l'administration de la recherche aux archives ouvertes

Pour finir sur une note positive, il faut mentionner comment ces systèmes pourraient créer des synergies vertueuses. Les chercheurs se plaignent régulièrement, en petit comité (comme dans [cet échange en réunion de labo rapporté sur une liste de diffusion](#)) ou [sur les réseaux sociaux](#), de la saisie chronophage des informations dans HAL. Or, [de son propre aveu](#), « [le projet] Conditor sera un facteur d'incitation au dépôt dans les archives ouvertes en limitant la saisie de méta-données ». Cela se vérifie partout dans le monde, où [les logiciels CRIS font tout ce qui est techniquement possible pour faciliter le dépôt](#). Par exemple, les utilisateurs du logiciel Elements de Symplectic [peuvent repérer dans la base de données d'articles en biomédecine PubMed](#) (ou plus exactement son miroir européen Europe PMC) leurs publications et les charger dans leur archive institutionnelle : tous les champs connus sont pré-remplis et la licence de l'éditeur est affichée pour aider à choisir la version (preprint, postprint) qui peut légalement être déposée. Qui plus est, dès qu'un éditeur assouplit sa politique le logiciel va proposer spontanément au chercheur de déposer ses anciens articles concernés par le changement de licence (voir capture ci-après) !

|                                  |   | Approved | Pending | Export |
|----------------------------------|---|----------|---------|--------|
| Books                            | Add new <a href="#">book</a>            | 0        | 0       |        |
| Chapters                         | Add new <a href="#">chapter</a>         | 0        | 0       |        |
| <a href="#">Conferences</a>      | Add new <a href="#">conference</a>      | 49       | 4       |        |
| <a href="#">Journal articles</a> | Add new <a href="#">journal article</a> | 63       | 6       |        |
| Patents                          | Add new <a href="#">patent</a>          | 0        | 0       |        |
| Reports                          | Add new <a href="#">report</a>          | 0        | 0       |        |
| Software                         | Add new <a href="#">software</a>        | 0        | 0       |        |
| <a href="#">Others</a>           |   | 0        | 0       |        |
| <a href="#">All</a>              |   | 112      | 10      |        |

Symplectic Elements alerte automatiquement l'auteur quand l'éditeur change sa politique et l'article peut être déposé. Source : John Fearnis, « Symplectic Elements and Repository Tools », RoMEO and CRIS in Practice, 1er avril 2011

Ainsi, en même temps qu'elle faciliterait l'administration de la recherche, l'adoption large de logiciels CRIS modernes pourrait faciliter l'accès libre aux publications scientifiques... qu'attend-on encore pour réagir ?

# Annuaire de chercheurs et valorisation de l'expertise des laboratoires

## Un prix Nobel en manque de visibilité

En octobre 2013, Martin Karplus recevait le prix Nobel de chimie en compagnie de deux autres chercheurs pour leurs travaux sur la modélisation multi-échelle de systèmes chimiques complexes. Et, petit cocorico, on apprenait que Martin Karplus travaille à la fois à l'université de Strasbourg et à l'université de Harvard. Or voici quelle **visibilité son laboratoire français donne à ce chercheur sur son site web** :



Profil de Martin Karplus dans l'annuaire de son laboratoire strasbourgeois

Il s'agit, on le rappelle, d'un chercheur réputé et admiré, partagé entre la France et Harvard ! Non pas qu'il faille lui consacrer un "hall of fame" (quoique, si on se donnait les moyens qu'il faut en communication scientifique...) mais on pourrait s'attendre à une présentation de son travail, ses réussites, ses projets, ses enseignements. Ce que **l'université de Harvard a parfaitement compris** comme le montre la capture ci-après.

Et plus encore si on clique sur le **lien vers son équipe de recherche**, avec une présentation longue de ses travaux et sa bibliographie complète...

Ce constat n'est pas nouveau : si les sites web des universités et laboratoires français pèchent souvent par leur habillage ou leur contenu, **la partie « annuaire » ou « profils » est sans doute encore plus mésestimée**. Julien Pierre, qui a mené une thèse en sciences de l'information et de la communication sur l'identité numérique, a toujours soulevé l'écart pouvant exister entre la page institutionnel du chercheur et ses autres présences en ligne (réseaux sociaux, apparitions médiatiques, page Wikipédia, blog etc.). Comme dans le **diaporama ci-après, présenté lors d'une journée de formation à l'URFIST de Lyon**, le 9 juin 2011 :

## Pourquoi est-ce important ?

Pour répondre à cette question, on pourrait arguer sous forme de boutade que si les Américains le font, c'est bien que l'investissement est justifié ! Mais voici quelques réponses argumentées.

### Le « marché » de l'enseignement supérieur et de la recherche est mondial

Les meilleurs professeurs passent d'une chaire à l'autre et d'un pays à l'autre, quand ils n'occupent pas plusieurs postes

The screenshot shows the Harvard University website profile for Martin Karplus. At the top, it says 'HARVARD UNIVERSITY' and 'FACULTY OF ARTS AND SCIENCES | HARVARD.EDU'. Below that is the 'Department of Chemistry & Chemical Biology' with a search bar. A navigation menu includes 'HOME', 'Academics', 'Research', 'People', 'Resources', 'News & Events', and 'About'. The profile for Martin Karplus is titled 'HOME / PEOPLE / Martin Karplus' and identifies him as 'Theodore William Richards Professor of Chemistry, Emeritus'. It features a portrait photo and a paragraph about his research in molecular structure and dynamics. To the right is a 'PEOPLE' list with categories like 'Faculty & Lecturers', 'Staff', and 'Researchers'. Below the photo are sections for 'CONTACT INFORMATION' (p: 617-495-4018), 'WEBSITES' (Karplus Research Group), 'CATEGORY' (Faculty & Lecturers), 'RESEARCH AREAS' (Theoretical), and 'RESEARCH GROUP AFFILIATIONS' (Karplus Group Members).

Profil de Martin Karplus dans l'annuaire de son université américaine

The slide is titled 'Pages Web institutionnelles' and contains a bullet point: 'Pas d'info, pas de discipline, pas de champ de recherche, pas de publi'. Below the text is a screenshot of a website with a complex, cluttered layout. The website header includes logos for 'UNIVERSITÉ BORDEAUX I Sciences Technologiques' and 'UNIVERSITÉ BORDEAUX II Nouvelles Sciences'. The main content area is filled with numerous small, overlapping links and text elements, making it difficult to navigate. The slide footer includes the date '07/02/2011', the name 'Julien PIERRE - LRFIST Nancy', and the number '13'.

Les pages web institutionnelles des chercheurs vues par Julien Pierre

en même temps (comme Martin Karplus). Prendre soin d'un chercheur de dimension internationale en lui donnant la visibilité qu'il mérite devrait aller de soi ! Et comment faire connaître au monde entier qu'il travaille dans votre laboratoire plutôt que chez le voisin si vous ne l'affichez pas clairement ? Encore plus que l'affiliation des articles qu'il ou elle signera, le site web reste votre premier canal de communication et c'est un investissement efficace pour attirer plus d'étudiants, de candidats, de collaborateurs... et de financements venus du monde entier.

## Les compétences des chercheurs sont uniques

Un des concepteurs du [portail de la recherche en SHS Isidore](#) nous confiait que les journalistes ont très vite compris l'intérêt de ce moteur de recherche pour repérer des experts à interroger : sur le conflit israëlo-palestinien, sur les rythmes scolaires, sur la neutralité du net, sur l'enfouissement des déchets nucléaires... La richesse des laboratoires, et plus largement de la recherche française, repose sur ces femmes et ces hommes qui ont développé des expertises très spécifiques et des compétences extrêmement pointues. Cette ressource principale, la matière grise des labos, doit être mise en avant comme telle. Car personne ne peut prédire qui aura besoin de votre expertise un jour !

## L'information doit être accessible en un clic

Le retour d'expérience d'Isidore comportait un bémol : une fois parvenus à l'article scientifique qui les intéresse, les journalistes ne savent pas toujours comment contacter le chercheur en question — qu'il ait changé d'affiliation, que ses coordonnées ne soient pas publiées, qu'il ait pris sa retraite etc. Là où un annuaire bien indexé dans Google prendrait tout son sens ! L'exemple des journalistes peut se transposer au monde socio-économique (petites ou grandes entreprises, associations, collectivités, établissements publics...) quand il a besoin d'un expert pour une mission de conseil, pour discuter d'une nouvelle technologie, etc. Et même aux chercheurs qui voudraient contacter des collègues d'autres disciplines pour lancer des collaborations transversales ! Ainsi, le chercheur en informatique David Monniaux [s'est déjà plaint sur son blog influent](#) d'avoir eu « les pires difficultés à trouver les coordonnées où joindre des collègues de sciences humaines (ni page Web, ni email indiqué, ni annuaire...) ».

## Si vous ne le faites pas... personne ne le fera à votre place !

Le Commissariat à l'énergie atomique et aux énergies alternative (CEA), qui n'est pas le dernier pour la valorisation de la recherche et les collaborations industrielles (notamment en raison de son statut d'établissements public à caractère industriel et commercial), a pris ce problème au sérieux en lançant la base COLA (Compétences des laboratoires). Celle-ci rassemble les données fournies par les unités de recherche concernant leurs domaines de compétences, axes de recherches et collaborations. Accessible à tout agent du CEA, [elle est utilisée notamment par les conseillers chargés de la diffusion technologique en régions](#) pour identifier des interlocuteurs en réponse à des demandes d'industriels.

En réalité notre titre est exagéré et certains services font déjà ce travail à votre place : [Expervnova vend son moteur de recherche](#) aux industriels pour leur permettre d'identifier le bon expert, ou d'obtenir un panorama des compétences, acteurs et travaux en lien avec leur problématique. On pense également aux réseaux sociaux de chercheurs comme [ResearchGate](#) ou [Academia.edu](#), qui échappent aux établissements de tutelle. Or vous ne voudriez pas laisser à un tiers le monopole de la mise en relation avec vos experts, n'est-ce pas ?

## Les annuaires d'université sont morts, vive les Research Networking Systems !

En 2009, le National Center for Research Resources affilié aux NIH américains [a lancé un appel à projets](#) pour concevoir une infrastructure permettant de connecter les personnes et les ressources (données, modèles animaux, réactifs, matériel, outils...) afin de **faciliter la découverte d'individus et de ressources scientifiques par les chercheurs et étudiants, pour encourager les collaborations interdisciplinaires et les échanges scientifiques**. Ce fut le point de départ du premier Research Networking System (RNS), ou système de réseautage pour la recherche, [pilote par l'université de Floride avec une dotation de 12 millions de dollars](#). Il s'agit de VIVO, d'après une technologie développée à l'université de Cornell depuis 2003, sur lequel nous reviendrons.

La spécificité des RNS, même s'ils peuvent remplir d'autres fonctions comme la gestion d'un portefeuille de projets de recherche, est qu'ils sont basés sur les profils des individus. Et puisqu'ils visent à faciliter les collaborations, ils donnent des éléments souvent déterminant pour le succès ou de l'échec d'une collaboration : besoins et objectifs des chercheurs, caractéristiques du projet, politique des tutelles, normes disciplinaires, contraintes institutionnelles...

## Les annuaires de chercheurs au crible de 8 bonnes pratiques

Les expérimentations en matière d'annuaires d'universités ne manquent pas, heureusement ! Mais comment juger de ce qui relève des bonnes ou des mauvaises pratiques ? Titus Schleyer de l'université de Pittsburgh et ses collaborateurs [ont mené l'enquête dans le champ biomédical](#), en mêlant un [brainstorming KJ](#), des entretiens avec des chercheurs, une analyse de la littérature et des observations en contexte. Ils ont conclu qu'il existait 8 pré-requis à un bon RNS. Ces 8 pré-requis forment les chapitres suivants, que nous tenterons à chaque fois d'illustrer avec des solutions existantes.

## 1. Bon rapport coût/bénéfice de la création et la mise à jour des profils en ligne

Selon une étude britannique, la mise à jour des pages web des chercheurs, laboratoires, départements... d'une université représenterait 10 heures de travail par page et par an avec un outil de publication web classique. Pour un établissement de 1 000 chercheurs, organisé en 5 départements et 30 laboratoires, cela revient à 1 035 pages x 10 h/an soit 10 350 h/an, coûtant 310 500 € / an à raison de 30 € l'heure chargée d'ingénieur d'étude. Il semble bien plus efficace de collecter et mettre en forme automatiquement les informations disponibles dans les laboratoires ou les bases de données tierces (brevets, publications, dotations...). Or justement, les laboratoires utilisent de plus en plus ces logiciels de gestion des informations tout au long du cycle de vie de la recherche, que nous avons appelés Current Research Information Systems (CRIS) au chapitre précédent. Nous estimions alors le coût de déploiement d'un logiciel CRIS dans un établissement entre 12 000 et 25 000 € par an, bien inférieur aux 310 500 € de mise à jour manuelle.

Le choix du CRIS (ou à tout le moins d'un outil d'annuaire basé sur un CRIS) est donc un choix économique, mais c'est aussi un choix humain : il faut éviter les redondances de saisies d'information déjà nombreuses tout au long du cycle de vie de la recherche et chercher plutôt à exploiter au mieux l'information disponible.

En plus de diminuer les coûts, on peut chercher à augmenter les bénéfices. Pourquoi ne pas envisager, par exemple, que les pages de profils créées dans l'annuaire puissent être utilisées pour générer à la volée le CV des chercheurs ? En utilisant (tant qu'à faire) l'un des nombreux formats de CV qui se standardisent à travers le monde, à la demande notamment des institutions et des financeurs de la recherche :

- **Europass CV** est un standard européen proposé par le Forum européen sur la transparence des qualifications professionnelles, lancé en 1998 en vue de réunir les partenaires sociaux et les autorités nationales en matière d'enseignement et de formation professionnels autour du thème de la transparence des qualifications ;
- **le CV commun canadien** permet aux chercheurs de transmettre de façon standardisée les informations requises par un réseau d'organismes de financement fédéraux, provinciaux et de recherche à but non lucratif ;
- **NIH Biosketch** est un formulaire biographique requis par les NIH dans les réponses à leurs appels à projets de recherche ;
- **SciENCv** est un projet de plusieurs agences fédérales américaines pour créer un format standard de CV en ligne à l'usage des chercheurs, à partir duquel pourra notamment être généré une note biographique NIH Biosketch.

Ces exemples de formats n'ont pas été donnés au hasard : ce sont les formats d'export que l'on trouve dans le **logiciel Converis** (v. 5.4), un CRIS commercialisé par Thomson Reuters. La capture ci-après montre un **exemple de page de profil créée avec Converis**.

L'export vers l'un ou l'autre de ces formats de CV se retrouve dans d'autres logiciels sur lesquels nous reviendrons (VIVO, Digital Vita, UNIWeb, Elements, Pure...). Signalons également que le groupement euroCRIS **travaille actuellement** à aligner son schéma de données standard CERIF (qui sous-tend la majorité des CRIS) avec le **Currículum Vitae Normalizado español**.

## 2. Représentation fidèle des chercheurs grâce à une information riche, complète et à jour

Plus les profils seront complets, plus ils seront utiles pour déterminer l'opportunité d'une collaboration et faciliter la prise de contact. On peut s'attendre à trouver en particulier les informations suivantes :

- les accréditations (titres, diplômes, certificats, distinctions...) certifiant de la compétence générale dans tel ou tel domaine : la médecine, le droit, les sciences de gestion, etc. ;
- les expériences professionnelles (projets ou postes occupés, qui se traduisent par des productions type publications, brevets, modules de cours...) précisant les thèmes sur lesquels portent l'expertise du chercheur ;
- les champs d'intérêt déclarés fournissant des informations sur les motivations actuelles du chercheur ;
- les coordonnées permettant de savoir comment joindre le chercheur.

Toutes ces informations sont utiles à la personne qui recherche un profil précis.

**En pratique, c'est souvent la liste de publications qui est mise en avant**, malgré sa limite évidente (elle est rétrospective et ne dit rien des intérêts actuels ou futurs du chercheur). C'est pourquoi de nombreuses solutions d'annuaire s'adosent à l'archive ouverte institutionnelle. En effet, celle-ci recense la production de l'établissement et il sem-

The screenshot shows the University of Stirling website. The main navigation bar includes links for Home, Courses, International, Campus life, Research, Services we offer, Alumni & supporters, and About us. The 'Research' section is highlighted. Below the navigation, there is a search bar and a list of research areas including Sport, Health & Exercise, Health & Well-being, Education, Aquaculture, Culture and Heritage, Media, Stirling Media Research Institute, Public Relations and Journalism, Film and Media, Social Policy, Environment, Computing, Science and Maths related, Business and Economy, Psychology, Law, and Philosophy. The profile for Prof Gerard Hastings is displayed, featuring a photo, contact information (Email: gerard.hastings@stir.ac.uk, Phone: +44 (0)1706 467260, Fax: +44 (0)1706 467745), and a detailed 'About' section. The 'About' section describes his academic background, his role as a Special Advisor to the House of Commons Health Select Committee, and his work with the WHO. It also lists his research interests in critical and social marketing research and his employment history as Professor at the Institute for Social Marketing and Researcher in Socio-Management. There are also tabs for Projects (14), Publications (129), and Tags (6).

Exemple de profil dans l'annuaire de l'université de Stirling généré par Converis

ble logique de lui ajouter une couche « annuaire ». Ainsi, l'entrepôt Okina de l'université d'Angers, dont le lancement public est **prévu à la fin de l'année 2014**, a particulièrement soigné l'entrée par les laboratoires et les chercheurs de l'établissement (voir captures ci-après).

Okina est un développement « maison » basé sur Drupal 7. D'autres solutions clés en mains existent, comme **DSpace-CRIS** qui est une surcouche CRIS au fameux logiciel libre d'archive ouverte DSpace. Développé par la société italienne Cineca sous une licence libre, DSpace-CRIS est utilisé notamment par l'**entrepôt institutionnel de l'université de Hong Kong** (voir capture ci-après).

De son côté, le **logiciel Elements** commercialisé par Symplectic en Grande-Bretagne peut se brancher à un entrepôt DSpace, Fedora, ePrints ou IntraLibrary.

De la même façon, **il peut être intéressant de collecter l'information disponible ailleurs (typiquement dans des bases bibliographiques), en utilisant si possible les avantages du web sémantique pour donner tout son sens aux données récoltées et faciliter le dédoublenage.** C'est l'**approche suivie** par le **portail lyonnais SHSdocNET** qui « met en œuvre des technologies innovantes du web social et sémantique » :

- moissonnage de données (recherche automatique) à partir de toutes les sources d'information et de documentation publiques telles que **Isidore**, **HAL-SHS**, **SUDOC**, etc. ;
- sémantique : technologie s'appuyant sur les ontologies généralistes comme Rameau et spécialisées comme

### Institut des Sciences et Technologies Moléculaires d'Angers

**Site internet:** <http://moltech-anjou.univ-angers.fr/>  
**Code:** UMR CNRS 6200  
**Numéro national:** 200411667U  
**Tutelle(s):** Université d'Angers  
 CNRS  
**Adresse:** Université d'Angers - UFR Sciences, 2 bd Lavoisier, 49045 Angers  
**Pays:** fr  
**Année de création:** 2004  
**Pôle:** Matériaux

#### Publications

Consulter la liste des publications de ce laboratoire

#### Directeur

Marc Sallé

#### Membres

Consulter la liste des membres de ce laboratoire

#### Dernières publications

VAJPAYEE V., BIVAUD S., GOEB S., CROUÉ V., ALLAIN M., POPP B. V., GARCI A., THERRIEN B., SALLÉ M. « **Electron-Rich Arene–Ruthenium Metalla-architectures Incorporating Tetrapyrrolyl–Tetrathiafulvene Donor Moieties.** ». *Organometallics*. 2014. Vol. 33, p. 1651 - 1658.

CAUCHY T. « **Transfert de charge et effet d'antenne par TD-DFT. Une rationalisation par différence de densité électronique.** ». Rennes : [s.n.], 2014.

GOEB S., BIVAUD S., CROUÉ V., VAJPAYEE V., ALLAIN M., SALLÉ M. « **A Self-Assembled Electro-Active MBL4 Cage Based on Tetrathiafulvalene Ligands.** ». *Materials*. 2014. Vol. 7, p. 611 - 622.

BIVAUD S., GOEB S., BALANDIER J. - Y., CHAS M., ALLAIN M., SALLÉ M. « **Self-Assembled Cages from the Electroactive Bis(pyrrolo)tetrathiafulvalene (BPTTF) Building Block.** ». *European Journal of Inorganic Chemistry*. 2014. Vol. 2014, p. 2440-2448.

CHOI J. - W., KIM C. - H., PISON J., OYEDELE A., TONDELIER D., LELIEGE A., KIRCHNER E., BLANCHARD P., RONCALI J., GEFFROY B. « **Exploiting the potential of**

#### Equipes

CIMI - Chimie Inorganique, Matériaux et Interfaces

ERDySS - Electrochimie, Réactivité, Dynamique et Structuration de Surfaces

MINOS - Molecular Interactions, Non linear Optics and Structuring

SCL - Systèmes Conjugués Linéaires

SOMaF - Synthèse Organique et Matériaux Fonctionnels

Aperçu de l'entrepôt institutionnel de l'université d'Angers

### Eric Levillain

Voir
Publications
Contact



**Nom:** Levillain  
**Prénom:** Eric  
**Laboratoire:** Institut des Sciences et Technologies Moléculaires d'Angers  
**Equipe:** Electrochimie, Réactivité, Dynamique et Structuration de Surfaces  
**Présentation:**

**Fonction :**

- Directeur de Recherche (DR1 CNRS)

**Adresse professionnelle :**

- Laboratoire MOLTECH Anjou / UMR 6200 CNRS (Site internet)
- Université d'Angers
- 2, bd Lavoisier
- 49045 ANGERS Cedex - FRANCE
- Tél. : 02 41 73 50 95
- Fax : 02 41 73 54 05
- Email : [eric.levillain@univ-angers.fr](mailto:eric.levillain@univ-angers.fr)

**Disciplines enseignées**

- Electrochimie
- Traitement de données

**Thèmes de recherche**

Equipe : ERDySS (Site internet)

Activités centrées sur :

- La mise en évidence des relations structure/propriétés de précurseurs d'objets moléculaires en utilisant les outils théoriques

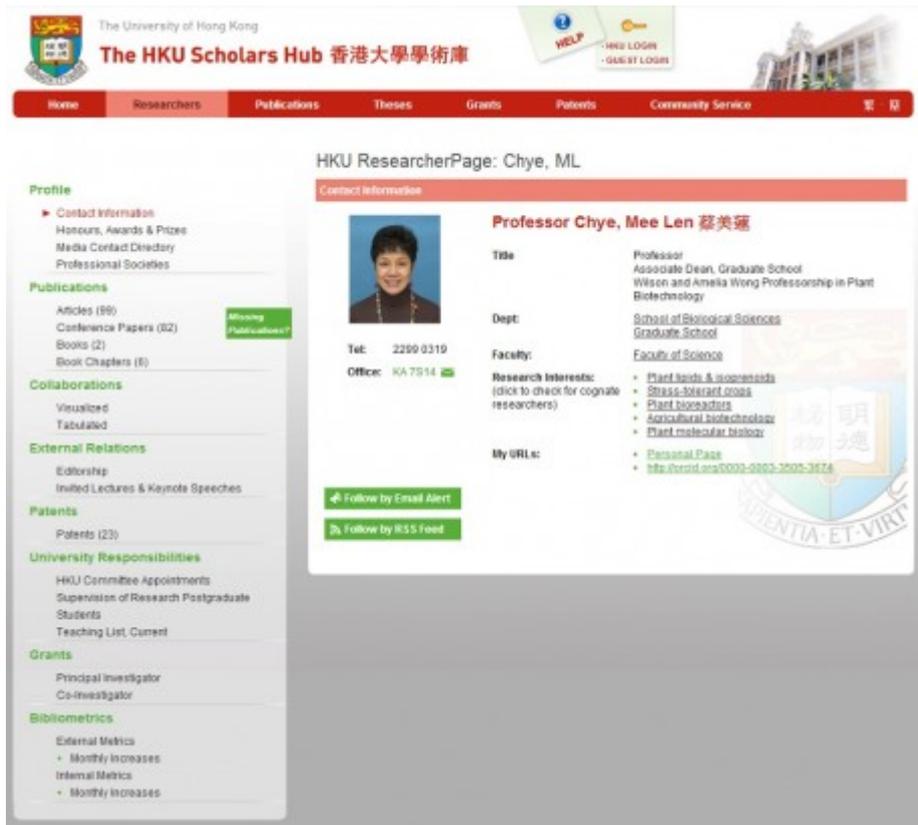
Aperçu de l'entrepôt institutionnel de l'université d'Angers

celle en cours d'élaboration à l'Institut des sciences de l'Homme, permettant d'améliorer la recherche de compétences et de proposer des compétences proches lors des recherches ;

- consolidation des informations par les individus référencés et/ou les gestionnaires des laboratoires auxquels ils appartiennent.

### 3. Exploitation des réseaux sociaux

**Au moment d'entrer en contact avec quelqu'un, vous aurez plus de chances de succès si vous passez par un contact commun ;** et si vous cherchez une expertise éloignée de vos centres d'intérêt, il est souvent utile de solliciter une connaissance dont vous pensez qu'elle saura vous faire des recommandations (par exemple un directeur ou

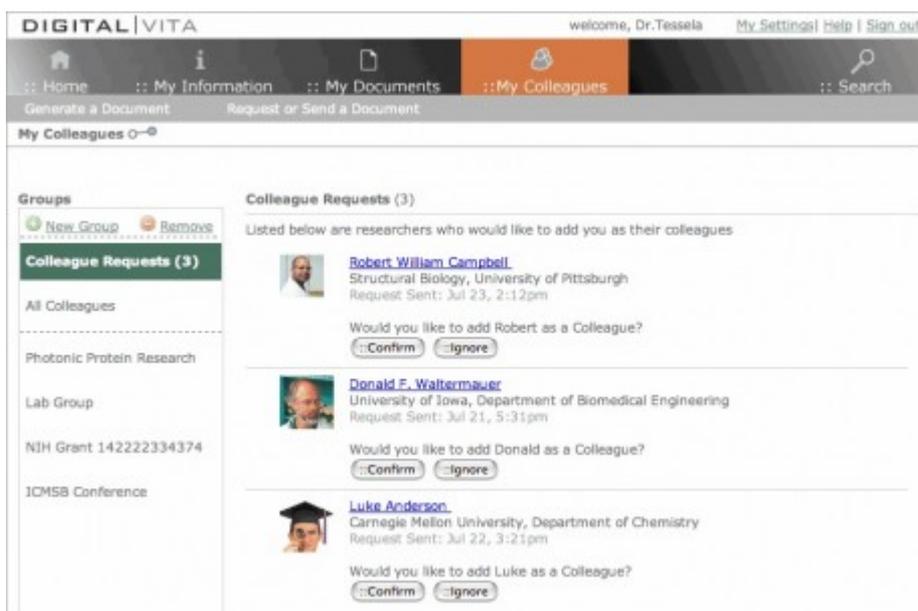


Aperçu de l'entrepôt institutionnel de l'université de Hong Kong

responsable, au réseau plus étendu). Par conséquent, il est important de pouvoir exploiter les liens existant entre personnes : co-auteurs, collègues du même établissement ou même département, etc.

Digital Vita et Profiles RNS proposent une double approche pour répondre à cette question :

- l'appartenance à un même département de recherche et les co-publications sont implicitement comptabilisées comme liens de proximité (c'est le « réseau passif ») ;
- les chercheurs peuvent déclarer explicitement leurs contacts à travers des « colleague requests » calqués sur les « friend requests » de Facebook (c'est le « réseau actif ») — voir capture ci-après.



Exemple de « Colleague Requests » dans Digital Vita.

#### 4. Evaluation des savoir-être des collaborateurs potentiels (personnalité, style de travail...)

L'affinité thématique ne suffit pas à déterminer si une collaboration peut fonctionner ou non. On peut par exemple imaginer un outil qui évaluerait également la compatibilité des styles de travail, la disponibilité des chercheurs et leur envie de démarrer une nouvelle collaboration. Ces informations sont plus difficilement objectivables et disponibles, et nous ne les avons retrouvées dans aucun des outils testés. A défaut, il est possible de s'appuyer sur l'existence de contacts communs (cf. chapitre 3.) pour former une première opinion sur la personne.

#### 5. Utilisation d'indicateurs multiples sur les activités de collaboration passées

Des études ont montré que la liste des co-auteurs ne donne qu'une image imparfaite du réseau de collaboration d'un chercheur (Bordons et Gómez, 2000 ; Katz et Martin, 1997). Par conséquent, il convient de la compléter avec d'autres données.

Or il arrive souvent que dans un projet de recherche pluridisciplinaire, tous les chercheurs aient une vision globale de l'ensemble mais ne collaborent spécifiquement qu'avec les chercheurs du même « *work package* », qui seront leurs co-auteurs. Dans ce cas, il peut être intéressant de relier entre eux tous les chercheurs qui ont participé, à quelque titre que ce soit, à un même projet. Et justement, les CRIS alimentés par l'université contiennent les informations relatives aux projets financés, à défaut de pouvoir les obtenir auprès des financeurs eux-même comme nous l'avons constaté dans le premier chapitre de notre enquête (voir capture ci-après).

Profil de l'annuaire de l'université du Nebraska généré par Pure, faisant apparaître les bourses de financement

Par ailleurs, **il peut s'avérer intéressant de distinguer dans le profil du chercheur ce qui relève de ses intérêts passés et ce qui relève de ses intérêts présents.** Digital Vita permet justement de qualifier chaque domaine d'intérêt selon la typologie suivante :

- Current Major ;
- Current Minor ;
- Non-current Major ;
- Non-current Minor.

Plus avancé encore, le logiciel **Profiles RNS** utilisé notamment par le **Clinical & Translational Science Institute** rattaché à l'université de Californie San Francisco a choisi d'illustrer l'évolution au fil du temps des centres d'intérêt de ses chercheurs sous la forme d'un graphique très parlant (voir capture ci-après).

**Titus KL Schleyer, DMD, PhD**

**Primary Appointment:**  
Clem McDonald Professor of Biomedical Informatics and Director, Regenrief Institute, Inc., Center for Biomedical Informatics

**Peer-reviewed Publications:**  
88

**Biography**  
Titus Schleyer, DMD, PhD, is Clem McDonald Professor of Biomedical Informatics at Indiana University and the Director of the Center for Biomedical Informatics at the Regenrief Institute (<http://www.regenrief.org/>). He holds DMD degrees from the University of Frankfurt, Germany, and Temple University, Philadelphia, as well as a PhD degree in molecular biology from the University of Frankfurt and an MBA degree in health administration from Temple University. Dr. Schleyer has been active in informatics research since 1995, conducting seminal research on electronic health records, and Internet applications, workflow and human-computer interaction in dentistry. Dr. Schleyer's informatics research is primarily funded by the National Institutes of Health. He publishes regularly on informatics topics in major journals, such as the Journal of the American Dental Association and the Journal of the American Medical Informatics Association. He also serves as director of the Biomedical Informatics core of the Indiana Clinical and Translational Sciences Institute. In 2006, Dr. Schleyer was elected to the American College of Medical Informatics for his significant and sustained contributions to the field of biomedical informatics. More information about Dr. Schleyer is available at <http://about.jhu.edu/schleyer>.

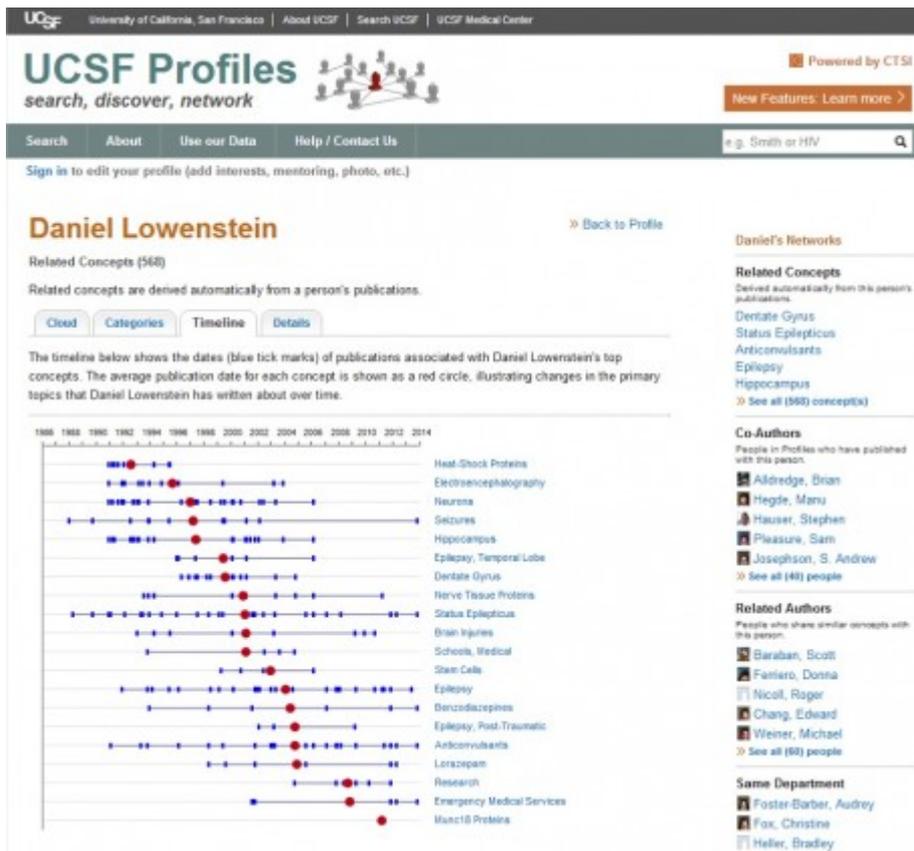
**Research/Scholarship Interests**  
algorithmic; American Dental Association; **affinity to computers**; Biomedical Research; biosentry; biostatistics; collaborator; **computer systems**; computer-assisted instruction; **computer-based patient records**; computer-supported cooperative work; **computers**; Consumer health information; Cooperative Behavior; **decision making**; **computer-assisted**; Dental Care for Chronically ill; **dental informatics**; **dentistry**; **education**; **dental**; Health education; humans; Information Management; **information science**; Information Services; **Internet**; **medical informatics**; **medical records systems**; **computerized**; Health Nomenclature; Pattern Recognition; Automated; Pilot Projects; Practice Guidelines as Topic; **practice management**; **dental**; Research Personnel; **software**; Speech Recognition Software; Systematized Nomenclature of Medicine; Terminology as Topic; Translational Medical Research; unified medical language system; Universities; user-computer interface; Incubation; Controlled

(Legend: **Current Major** / **Current Minor**, Non-current Major/ Non-current Minor)



**Contact Information**  
Regenrief Institute  
410 W 10th St,  
Indianapolis, IN 46202  
**Phone:** (317) 274-9204  
**Work:**  
[tschleyer@regenrief.org](mailto:tschleyer@regenrief.org)  
**Bookmark this Profile**  
[Print](#)

Présentation des domaines d'intérêt dans Digital Vita



**UCSF Profiles**  
search, discover, network

Powered by CTSI  
New Features. Learn more >

Search About Use our Data Help / Contact Us

Sign in to edit your profile (add interests, mentoring, photo, etc.)

**Daniel Lowenstein** [Back to Profile](#)

**Related Concepts (568)**  
Related concepts are derived automatically from a person's publications.

[Cloud](#) [Categories](#) [Timeline](#) [Details](#)

The timeline below shows the dates (blue tick marks) of publications associated with Daniel Lowenstein's top concepts. The average publication date for each concept is shown as a red circle, illustrating changes in the primary topics that Daniel Lowenstein has written about over time.

Timeline (1988-2014):

- Heat-Shock Proteins
- Electroencephalography
- Neurons
- Seizures
- Hippocampus
- Epilepsy, Temporal Lobe
- Dentate Gyrus
- Nerve Tissue Proteins
- Status Epilepticus
- Brain Injuries
- Schools, Medical
- Stem Cells
- Epilepsy
- Benzodiazepines
- Epilepsy, Post-Traumatic
- Anticonvulsants
- Lorazepam
- Research
- Emergency Medical Services
- Mucl18 Proteins

**Daniel's Networks**

**Related Concepts**  
Derived automatically from this person's publications.

- Dentate Gyrus
- Status Epilepticus
- Anticonvulsants
- Epilepsy
- Hippocampus
- [See all \(568\) concept\(s\)](#)

**Co-Authors**  
People in Profiles who have published with this person.

- Allredge, Brian
- Hegde, Manj
- Hausser, Stephen
- Pleasure, Sam
- Josephson, S. Andrew
- [See all \(48\) people](#)

**Related Authors**  
People who share similar concepts with this person.

- Baraban, Scott
- Fastiero, Donna
- Nicoll, Roger
- Chang, Edward
- Weiner, Michael
- [See all \(80\) people](#)

**Same Department**

- Foster-Barbar, Audrey
- Fox, Christine
- Helier, Bradley

Chronologie des centres d'intérêt d'un chercheur dans un annuaire Profiles RSN

Profils RNS sait également pondérer les résultats d'une recherche selon le rang du chercheur parmi les auteurs de la publication indexée avec le mot-clé, la fréquence de ce mot-clé dans les publications, l'ancienneté de la publication, etc. afin de retourner les résultats les plus pertinents et opérationnels possible !

## 6. Paramétrage de la confidentialité des informations

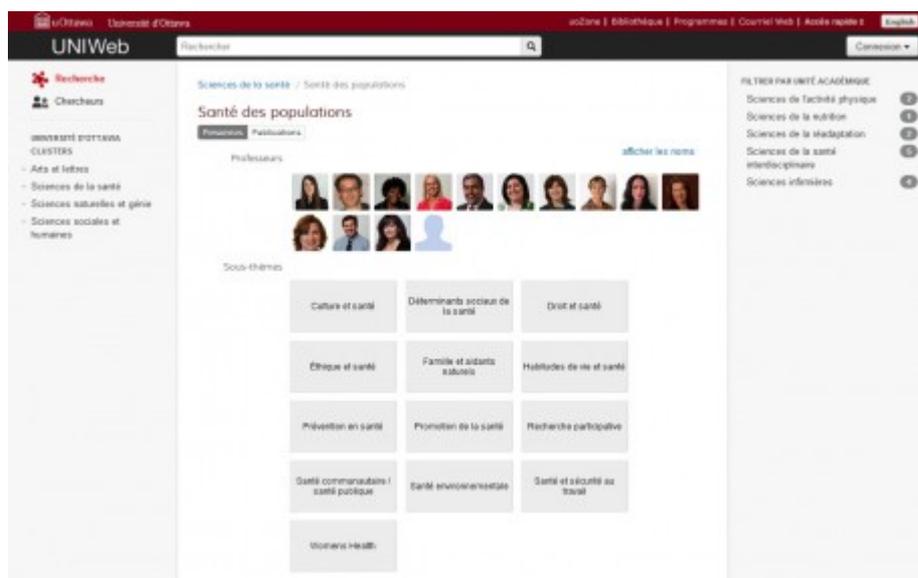
La protection de la vie privée et le droit à l'oubli sont devenus des exigences de notre vie numérique et toute entorse à ces règles se paierait cher : par conséquent, il n'est pas question de s'en passer. Qui plus est, un chercheur peut avoir plus ou moins envie d'être bien référencé dans un annuaire institutionnel, par exemple s'il reçoit déjà beaucoup de sollicitations (auquel cas on lui conseillerait plutôt de ne pas se cacher et d'en faire profiter les collègues !). Même s'il est de son devoir d'apparaître dans l'annuaire (en tant qu'agent public), il doit pouvoir maîtriser les informations publiées le concernant.

L'annuaire SHSdocNET assume par exemple une position dite « socio-éthique » qui nous semble intéressante. Celle-ci consiste à « refuser toute publicité, à ne jamais vendre d'extraits de la base de données, à ne jamais construire des statistiques nominatives, à garantir au mieux l'inviolabilité de la vie des chercheurs et à laisser le chercheur maître pour contrôler son information et décider par lui-même de ce qui est public et non ».

## 7. Recherche efficace à travers les disciplines

Une fois l'annuaire déployé, il ne s'avèrera utile que s'il permet effectivement aux chercheurs ou industriels de trouver les compétences dont ils ont besoin. Il faut un moteur de recherche, et un bon ! Plus que la recherche en texte libre, **l'utilisation de vocabulaires contrôlés comme le MeSH en santé et l'ACM Computing Classification System en informatique permettent de chercher les synonymes et termes connexes** (y compris les termes plus génériques ou spécifiques). Ça peut être utile quand on connaît mal le vocabulaire spécifique d'un domaine, pour chercher des compétences dans une autre discipline par exemple. Et l'utilisateur doit pouvoir affiner sa recherche avec autant de critères qu'il le souhaite.

Le logiciel **UNIWeb**, développé par la société canadienne **Proximify**, s'est fixé comme objectif de favoriser la recherche pluridisciplinaire en facilitant le rapprochement entre chercheurs d'une même institution. Pas question d'organiser l'information par faculté et département, **place à des mots-clés décrivant les thèmes et sous-thèmes de recherche** (même s'il reste toujours possible de filtrer sa requête par unité académique), qui forment des "grappes" de chercheurs ayant un domaine d'expertise en commun (voir capture ci-après).



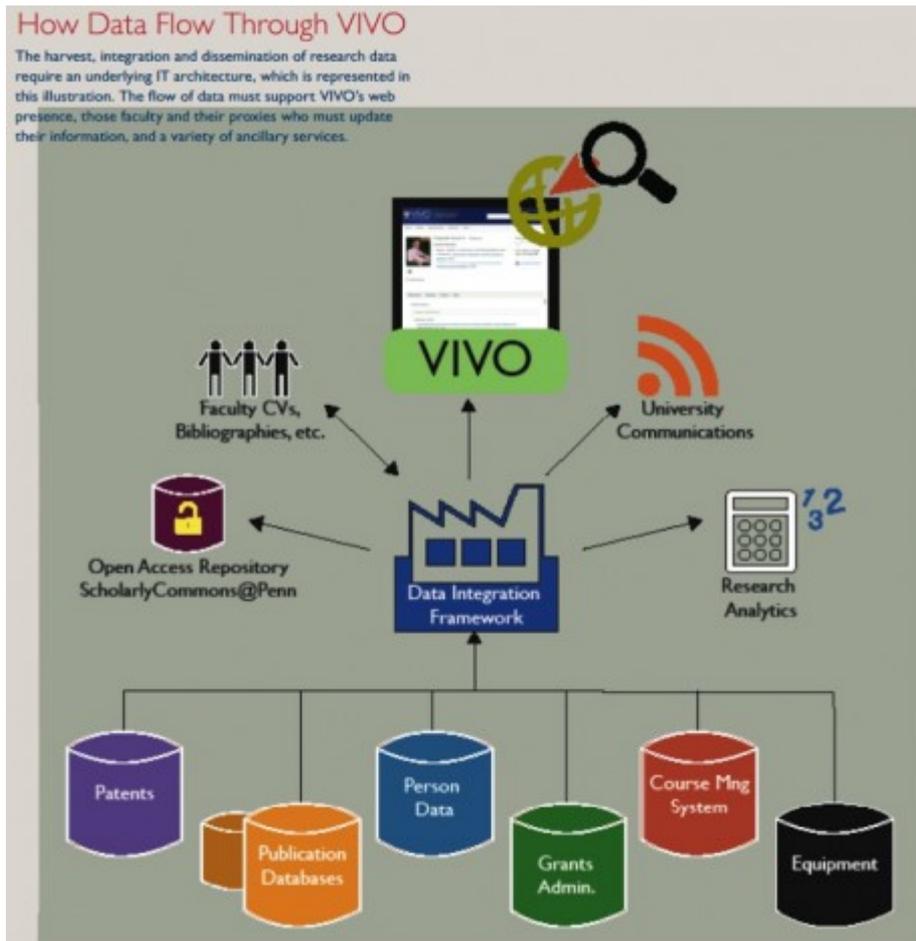
Page thématique « Santé des populations » dans l'annuaire UNIWeb de l'université d'Ottawa

Évidemment, plus votre annuaire embrassera tout le périmètre de votre établissement, plus il sera riche de compétences et de disciplines. De ce point de vue, **les projets d'annuaire menés à l'échelle d'un seul département sont forcément désavantagés... tandis que les initiatives d'interopérabilité entre annuaires sont récompensées** ! Deux exemples de méta-annuaires méritent d'être détaillés : VIVO et DIRECT2experts.

**VIVO est une application du web sémantique**, sous licence libre, capable de construire un annuaire de chercheurs à partir des données internes aux établissements et de bases externes comme **PubMed** (voir schéma ci-après). Bâti sur un schéma de données éprouvé et la technologie sémantique RDF, soutenu par une forte communauté de développeurs et d'utilisateurs (dont les plus grandes universités américaines comme Cornell) et appuyé par une subvention de 12,2 millions de dollars des instances fédérales américaines, c'est à la fois l'outil d'annuaire le plus ergonomique que nous ayons pu voir, et celui qui propose la navigation la plus riche (presque trop, voir par **exemple le profil du Pr. Nina Lauren Bassuk**).

Mais ce n'est pas tout : reconnu aujourd'hui comme un standard, VIVO est capable d'interopérer avec un tas de logiciels CRIS mais également avec d'autres instances de VIVO. D'où le puissant portail inter-institutionnel **VIVOsearch**, réunissant à titre expérimental 8 organismes de recherche : les universités de Cornell, d'Harvard, de l'Indiana et de Floride, les facultés de médecine de Ponce, St. Louis et Weill Cornell, et le Scripps Research Institute. Mentionnons également **VIVO Searchlight**, une extension pour navigateurs internet qui permet de requêter VIVOsearch à partir de n'importe quelle page web, par exemple pour trouver des experts relatifs à l'article que vous êtes en train de lire (voir capture ci-après).

Quant à **DIRECT2experts**, c'est un prototype de recherche fédérée qui a réussi en 6 mois à rendre interopérable 7 solutions logicielles utilisées dans 28 universités américaines. Conséquence de ces chantiers d'interopérabilité : des logiciels CRIS comme **Pure** ou **Elements** et des annuaires comme Profiles RNS permettent aujourd'hui d'être facilement présents dans les méta-annuaires VIVOsearch et/ou DIRECT2experts, avec la visibilité nationale afférente.



Intégration des données dans un annuaire VIVO

Exemple d'utilisation de VIVO Searchlight sur une page PubMed

## 8. Mise en évidence des connexions inattendues ou « non intuitives » entre chercheurs

Pourquoi ne pas envisager que les systèmes de réseautage pour la recherche vous aident à trouver des experts par sérendipité ou en proposant de façon pro-active des collaborateurs potentiellement intéressants ? Les algorithmes d'appariement (« *matching* ») sont très utilisés sur les sites de rencontre, alors pourquoi pas entre chercheurs également ? On trouve dans la littérature des algorithmes entraînés pour faire ce type de recommandations d'experts, en se basant **sur le réseau social des chercheurs** et/ou **sur les descripteurs MeSH de leurs publications**.

Parmi les logiciels du commerce, UNIWeb semble s'être penché sur la question et aide à révéler les connexions cachées

entre chercheurs : non pas leurs relations institutionnelles (co-auteurs ou membres du même laboratoire/département) mais leurs proximité thématique telle qu'elle est décrite par des mots-clés hiérarchisés. Il n'est pas encore question de calculer des scores de proximité mais ça donne déjà des **graphes de réseaux qui invitent à l'exploration** (voir capture ci-après).



Graphique de réseau d'un chercheur de l'université d'Ottawa

Elsevier va encore plus loin avec sa **technologie propriétaire Fingerprint Engine** incluse dans le logiciel CRIS Pure. Celle-ci fouille les publications associées aux chercheurs, les croise avec différents thésaurus afin d'en extraire les concepts clés, et crée l'empreinte du document qui permettra ensuite de « matcher » un chercheur avec les appels à projets pertinents, ou autre (voir schéma ci-après).

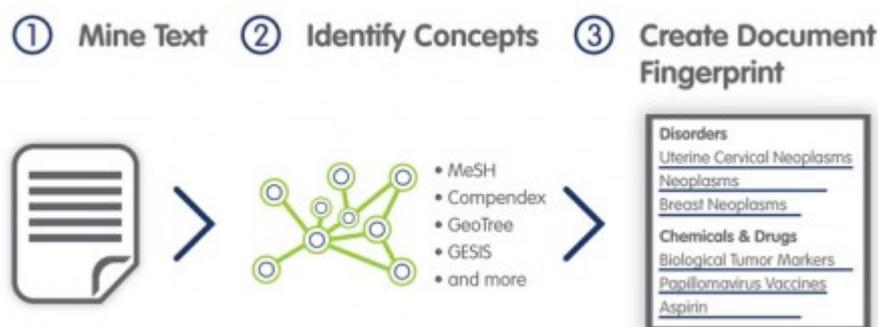


Schéma de principe de la technologie Fingerprint Engine d'Elsevier

## Conclusion : quatre solutions possibles

Pour conclure cet état des lieux, nous pouvons distinguer quatre approches permettant de bâtir l'annuaire des chercheurs d'une institution.

### 1. S'appuyer sur le service tiers ORCID

ORCID est un service offert par l'organisation à but non lucratif du même nom, qui permet à chaque chercheur d'obtenir son identifiant unique et d'alimenter la page de profil ORCID le concernant (voir capture ci-après). Chaque acteur du monde de la recherche est censé y trouver son intérêt :

- les chercheurs disposent d'un identifiant unique auquel ils peuvent rattacher toute leur production, sans risquer d'être confondus avec un homonyme (et **9 autres bénéfices listés ici**) ;
- les financeurs de la recherche peuvent obtenir en un clic le CV complet et formaté des porteurs de projet ;
- les éditeurs de revues peuvent déléguer à ORCID la gestion des métadonnées des auteurs (fin des soucis d'adresse électronique obsolète), trouver plus facilement les rapporteurs dont ils ont besoin, etc. ;
- les tutelles peuvent suivre précisément la production et l'impact de leurs chercheurs (y compris avec des alertes à chaque mise à jour), et échanger plus facilement avec d'autres institutions par exemple en cas de mobilité.

The image shows a screenshot of an ORCID iD profile for Marta Soler, M. The profile is organized into several sections:

- Header:** ORCID logo with the tagline "Connecting Research and Researchers". Navigation tabs include "FOR RESEARCHERS", "FOR ORGANIZATIONS", "ABOUT", "HELP", and "SIGN IN". Below these are links for "SIGN IN", "REGISTER FOR AN ORCID ID", and "LEARN MORE". A counter shows "838,549 ORCID iDs and counting. See more..."
- Profile Summary:** Name "Soler, M.", ORCID iD "http://orcid.org/0000-0001-6094-1288".
- Also known as:** Soler-Gallart, M.
- Country:** Spain
- Keywords:** Sociology, Language, Sociological Theory, Social impact
- Websites:** http://creasub.info/
- Other IDs:** Scopus Author ID: 36672084000
- Personal Information:**
  - Biography:** Marta Soler is Professor of Sociology at the University of Barcelona and current Director of CREA. She joined CREA as a junior researcher in 1992 and since then she participated in national and European studies on strategies to overcome exclusion. In 2001 she graduated from Harvard, with a dissertation on literary gatherings and social change (or the empowerment of the working class through breaking with cultural elitism). Some years later she spoke at the Haves Center at University of Wisconsin, Madison, about this study, with a lecture titled "Beyond Bourdieu". Back in Spain she continued this work elaborating on speech acts theory with the analysis of communicative acts, and applying this analysis to gender inequalities and gender violence. From this work highlights her book with the philosopher John Searle "Language and social sciences" (2004) and two studies funded by the Spanish Ministry of Science (Communicative acts and social inequality in gender relations; Impact of communicative acts in the construction of egalitarian masculinities). Besides, Marta Soler is the Academic Director of the Degree in Sociology at the University of Barcelona and Board member of the Catalan Sociology Association (Catalan Academy of Sciences).
- Education:** Harvard University (1996 to 2001), Doctor.
- Employment:** Universitat de Barcelona, Professor.
- Funding:**
  - IMPACT-EV. Evaluating the impact and outcomes of European SSH research: European Commission Directorate-General for Research and Innovation (2014 to 2017)
  - Impacto de los actos comunicativos en la construcción de nuevas masculinidades: Ministry of Sciences of Spain (2010 to 2012)

Exemple de profil ORCID

La « migration massive » de chercheurs d'une université voire d'un pays vers ORCID n'est pas irréaliste : fin 2013, la *Fundação para a Ciência e a Tecnologia* (FCT), principal financeur de la recherche au Portugal, a demandé à ses chercheurs de créer un profil ORCID et d'utiliser l'outil d'importation pour ajouter leurs publications depuis la base Scopus. En 3 semaines, 40 000 chercheurs portugais se sont enregistrés sur ORCID et ont importé plus de 75 000 références bibliographiques ! L'université de technologie Chalmers en Suède, l'université de Boston... s'y sont mis également, et un guide très bien fait est disponible en ligne pour passer à l'acte.

## 2. Déployer un logiciel CRIS

Nous l'avons vu, les logiciels CRIS que sont Pure, Elements ou Converis permettent également de construire un annuaire de chercheurs. L'annuaire constitue dans ce cas le sommet de l'iceberg, constitué par toutes les données produites dans l'université et gérées par le CRIS : dotations, publications, thèses, enseignements, collaborations, distinctions, brevets, impact... D'où un plus gros chantier à prévoir.

Par contre, il faut se garder des logiciels vieillissants utilisés en France comme GRAAL, Sangria ou Pleiade : ils sont loin du compte pour ce qui est d'offrir un annuaire digne de ce nom ! Ainsi, l'université Bordeaux Montaigne qui cherchait en 2012 à baser son annuaire de chercheurs sur Sangria, a finalement décidé de procéder autrement.

## 3. Installer une solution d'annuaire

Les logiciels VIVO, Profiles RNS ou UNIWeb sont spécialisés dans la gestion et la mise en ligne d'annuaires de chercheurs. Ils ont de nombreuses qualités, et les deux premiers ont l'avantage d'être sous licence libre et interopérables avec d'autres solutions d'administration de la recherche.

## 4. Adosser l'annuaire à un entrepôt institutionnel

Pour les établissements dont l'entrepôt institutionnel est déjà installé ou en projet, il peut être avantageux de bâtir par-dessus une « couche annuaire » qui pourra bénéficier de la richesse informationnelle de l'entrepôt. Et réciproquement, cela permet de placer l'archive ouverte au centre de la politique de l'établissement, avec la carotte (la visibilité, l'accessibilité) et le bâton (le mandat de dépôt, l'évaluation).

Le choix se situe alors entre un développement « maison » comme Okina à l'université d'Angers ou une solution clés en mains comme [DSpace-CRIS](#).

Enfin, il faut souligner à nouveau l'effort américain, poussé par quelques grandes universités et des facultés de médecine soucieuses de recherche translationnelle (« *from bench to bedside* »), de fédérer ces annuaires au sein de portails comme [VIVOsearch](#), [DIRECT2experts](#) ou [CTSAsearch](#).

Imaginez ce que l'équivalent en France pourrait faire à la visibilité des chercheurs et de la recherche française 😊

## Bonus

- notre [galerie Tumblr des sites web de labo les plus kitschs](#) du monde, alimentée par une douzaine de contributeurs à ce jour
- une [analyse stylistique du web académique de 1993](#), toujours vivace aujourd'hui

## COMMENTAIRES ET RÉACTIONS

Les prises de position que vous venez de lire dans ce livre blanc ont suscité un certain nombre de réactions. Voici quelques commentaires que nous avons sélectionnés pour la postérité...

**Yannick Maignien**, ingénieur de recherche au CNRS :

Excellent article et initiative. Je peux me tromper, mais je vous trouve optimiste sur la capacité de l'administration française à évoluer rapidement vers l'ouverture des données publiques de (ou plutôt sur) la Recherche.

Ayant été Directeur du TGE ADONIS (CNRS) et responsable de la création d'ISIDORE entre 2007 et 2011 (puis viré par le CNRS), sur ces mêmes finalités de transparence et d'interopérabilité, mais en SHS, je ne peux que témoigner de l'importance des obstacles de tout genre, de la part des directions des institutions de la Recherche ou de l'Enseignement supérieur...mais aussi de nombreux protagonistes qui font leur fond de commerce de l'Open ... (ce qui ne facilite pas les choses). Certains ont très bien compris que compétition, rivalités personnelles et carriérisme passaient par ce nouveau terrain de lutte !

Cf. un extrait : « Les nouvelles frontières numériques des sciences » <http://www.donneesdelarecherche.fr/spip.php?article186>

Cela dit, bravo pour votre étude et votre initiative, que je signe des deux bras.

**Stéphanie Dord-Crouslé**, chargée de recherche au LIRE (université Stendhal Grenoble 3 / université Jean Monnet Saint-Etienne / Université Lumière Lyon 2 / CNRS / ENS Lyon) :

Article très intéressant !

Pour info, j'ai porté un projet financé par l'ANR (Corpus 2007) et il a fallu que je me batte pour faire un rapport final (ils m'avaient administrativement « perdue » depuis 2 ans et semblaient n'y accorder aucune importance !) et pour savoir si je pouvais bien le mettre sur HAL !

Il y a en effet du boulot pour savoir qui est financé pour faire quoi et évaluer ce qui est effectivement produit au terme du financement...

**Gaïa**, blogueuse (sous pseudonyme) assidue sur l'enseignement supérieur et la recherche :

Puisque que vous m'y avez invité, voici quelques commentaires. La lecture du billet m'a beaucoup intéressé et me paraît contenir beaucoup d'éléments constructifs. D'autres sont plus discutables et je vais focaliser mes commentaires sur ces derniers.

L'article précise bien que les données auxquelles vous voulez accéder ne sont pas des données scientifiques mais des données sur la recherche (données administratives). En tant que chercheur (et membre d'université, labo, école doctorale) je peux vous assurer qu'on nous sollicite beaucoup pour fournir ce genre de données, avec des formats différents et toujours avec des questions posées un peu différemment ; souvent aussi avec des informations stupides et inutiles et qui prennent beaucoup de temps à agréger. S'il pouvait y avoir une amélioration du système, ça serait d'avoir un service unique de collecte de données dans lequel les financeurs, administrateurs et contribuables viendraient prendre ce qui les intéresse.

Vous voulez savoir « qui fait quoi » ou essayer de comprendre comment ça marche (la structure du système de la recherche). Je crois que vous vous faites de douces illusions. Mais je peux vous assurer que de l'intérieur personne n'y comprend rien (on fait juste illusion, pour ne pas paraître idiot ...). C'est devenu beaucoup trop compliqué. S'il y a là quelque chose à réclamer c'est d'abord un choc de simplification de grande ampleur.

Vous avez l'air de penser que pour avoir des données quantitatives sur la recherche il suffirait que les chercheurs mettent leurs travaux sur un serveur « archives ouvertes ». Je n'ai rien contre cette idée, au contraire, mais vous n'allez pas en extraire grand-chose, mis à part de compter les publications (en termes de données « administratives » sur la recherche).

Non les données quantitatives sur la recherche ne sont pas enfouies dans les organismes. Hormis les aspects financiers, la capitalisation des productions existent dans les moteurs de recherche (payants, de type Scopus ou Thompson Reuters). C'est là que vous trouverez beaucoup de choses sur la recherche française (et aussi partout ailleurs) ; selon les disciplines, les labos, les universités, les pays ... Certaines sont gratuites (exemple : Scimago). Comme mentionné un peu plus haut dans les commentaires, l'OST est aussi un formidable outil, avec beaucoup de données (certes ce n'est que macroscopique, pas possible d'aller voir le projet bidule financé par l'agence truc).

En filigrane, vous semblez reprocher à l'ANR des manquements. D'accord, critiquer l'ANR c'est de bon ton. Mais je pense que vous vous trompez de cible. L'ANR a fait un réel effort de transparence et d'incitation à ce que les projets et résultats soit disponibles et accessibles (obligation d'avoir un site web pour chaque projet, de rendre des compte périodiquement sur les résultats, etc ...). Si le travail est mal fait c'est certainement un peu à cause de l'agence qui n'arrive pas à bien tout suivre mais c'est aussi à cause des chercheurs qui font rarement ce travail consciencieusement. Par ailleurs l'ANR n'est que petite partie de la recherche. Vous n'aurez qu'une vision très parcellaire de la recherche avec l'ANR.

A mon sens la seule structure qui pourrait être adaptée à votre requête c'est l'AERES (agence d'évaluation). Cette agence a évalué tout le monde et a donc des données qualitatives et quantitatives sur tous les volets (financiers, production, effectifs, ...), à une granulométrie qui va jusqu'à l'équipe de recherche (quelques personnes). C'est une mine formidable d'informations. Malheureusement on vient de la fermer ...

Voilà pour ces quelques commentaires « critiques », qui ne doivent en aucun cas occulter la qualité de votre travail très intéressant.

**Emilien Schultz**, doctorant au GEMASS (université de Paris-Sorbonne / CNRS) sur les « effets de la création de l'Agence Nationale de la Recherche » :

Cet article est très intéressant, riche et effectivement souligne la difficulté que tout observateur peut avoir quand il s'intéresse à la recherche française : obtenir une vision d'ensemble. Et effectivement, l'accès aux données quand il est possible est réservé à un petit nombre de privilégiés (en fait, pour le cas abordé de l'ANR, cela est très fortement lié à la jeunesse de l'agence – créée en 2005 – qui n'a intégré l'ensemble de ses programmes que sur le tard, et que la critique nourrie subie d'une part de la communauté scientifique française n'a pas favorisé un projet de normalisation des pratiques de financement).

La transparence de l'utilisation de l'argent public, à partir du moment où il ne contribue pas à exposer trop facilement des chercheurs au feu de l'actualité, ne peut qu'être revendiquée. Je souhaiterais juste ouvrir un peu la discussion sur le risque de réifier l'objet "projet" auquel on pourrait attribuer une signification unique (sans pour autant prétendre ajouter des éléments neufs au débat, simplement rappeler quelques points qui semblent pertinents, car si la discussion sur les dérives de la scientométrie a déjà été fortement nourrie concernant la bibliographie, cela est moins le cas sur la question des financements). A moins de généraliser la feuille de temps et un suivi précis de la répartition des sommes sur l'ensemble des partenaires d'un projet (car souvent, il y a plus d'un laboratoire), on risque de créer une fiction au moment d'agréger les données. Ajouté à cela la balkanisation croissante des sources de financements (si on en reste à la recherche fondamentale qui ne bénéficie ni de fondations biomédicales, ni d'entreprises privées, la gamme de financement peut aller des appels internes aux organismes aux grands financeurs internationaux type HFSP en passant par les bourses de thèses ministérielles), l'ANR n'est ni négligeable, ni représentative dans beaucoup de situations. Sans même aborder le fait que le rapport à l'argent des différentes thématiques de recherche varie dans des proportions telles qu'à somme comparable, on compare des réalités radicalement différentes, et que des arrangements locaux plus ou moins importants contribuent à rendre flou la portée de telles données (entre les laboratoires qui mettent en commun tous leurs financements sur projets, et les laboratoires hautement compétitif où seule l'équipe est pertinente, comment identifier le niveau pertinent d'analyse). Et ça, c'est sans parler du problème de mesure d'impact, et de l'attribution de telle ou telle publication à un financement, ce qui pour certains domaines de recherche peut relever de la

fiction pure et simple. A trop rationaliser, on risque de forcer les éléments à rentrer dans les cases des tableaux excel, dont j'attends avec impatience l'analyse dans le prochain volet de l'enquête

Tout ça pour dire la trivialité suivante : *evidence based policy* menée à partir de constitution de base de données de ce type risque, à moins d'une vigilance constante sur les formes réelles d'interdépendance qui existent entre financements et pratiques de recherche, d'approfondir la dynamique de découplage déjà soulignée dans la tendance à la managerialisation de la recherche (cf. C. Vilkas (2009), « Des pairs aux experts: l'émergence d'un 'nouveau management' de la recherche scientifique ? », *Cahiers internationaux de sociologie*, (1), 61-79)

Encore une fois, merci pour cette enquête qui soulève de très nombreuses questions.

Un lecteur anonyme réagissant sur le site du *Monde* :

Pourquoi se limiter aux EPA, donc à vocation administrative, sachant que l'essentiel de la recherche hors universités est réalisée dans les EPST (Etablissements publics scientifiques et techniques) ? Ceux-ci font des bilans, annuels et pluriannuels, plus ou moins quantifiés. Les Régions en France ont aussi un rôle dans le financement de la recherche.

Gilles Saint-Martin :

Encore du temps consacré par les chercheurs à entrer des données dans ses systèmes administratifs au détriment du temps dédié à la recherche. L'analyse des publications des UR suffit à donner des informations sur l'orientation de la recherche en France. L'OST fait ça très bien. Ces informations sont accessibles sur le *Web of Knowledge*. Elles permettent un pilotage de la politique scientifique française.

En revanche, je souscris aux propos de l'article sur la nécessité « de faire advenir une politique de la recherche fondée sur des preuves (*evidence based*) ». L'utilité sociale des résultats de recherche analysée à travers les changements induits par ces résultats (l'impact) reste une dimension totalement absente du pilotage de la recherche publique en France.

Un lecteur fâché écrivant à la rédaction du *Monde* :

Je viens de prendre connaissance de la tribune intitulée « pour un accès aux données sur la recherche française ».

Depuis peu à la retraite après une carrière de 35 ans de délégué régional d'organisme de recherche d'abord à l'INSERM puis à partir au CNRS et passionné par le rôle et l'organisation de l'ESR en France, je suis désagréablement surpris par le contenu de cette tribune.

En effet ses rédacteurs méconnaissent de nombreuses sources d'information dont l'essentielle est l'Observatoire des Sciences et des Techniques (OST) GIP interministériel et interorganismes qui publie tout les deux ans un état de la recherche en France y compris la valorisation par régions, sites, disciplines, sous-disciplines sans oublier les rapports à la demande.

Ces rapports constituent une mine d'information et en plus ils sont téléchargeables et gratuits.

Les laboratoires quant à eux sont soumis à évaluation quadriennale par leur tutelle en « partenariat » avec l'AERES créée en 2006 mais effectivement leur contenu reste relativement confidentiel.

L'ANR est une agence de moyens permettant un pilotage politique mais ne peut constituer en rien une source plénière d'informations.

Je suis en tout cas surpris qu'un docteur en biologie et un ingénieur en agronomie se lancent sur ce créneau sans beaucoup de préparation semble-t-il.

Nicolas Alarcon, responsable du secteur documentaire Lettres-Sciences humaines à l'université de La Réunion :

Merci pour cet article édifiant. Je n'avais jamais entendu parler de CERIF, n'y a-t-il pas une certaine méfiance des acteurs français à utiliser des spécifications rédigées par (pour ?) la Grande-Bretagne, même si produites dans un cadre européen. Pourtant, on voit clairement les avantages pour l'administration de la recherche.

Sur l'intéropérabilité de GRAAL avec HAL, je serai moins enthousiaste que vous, je n'ai jamais vu d'exploitation massive de cette intéropérabilité. Sur le papier, ça semble fonctionner malgré les problèmes répertoriés mais avez-vous des exemples concrets de mise en production ?

Les retards de Caplab sont aussi très inquiétants mais le plus dur à avaler est que la solution proposée risque de ne même pas répondre aux standards en vigueur. C'est décourageant.

**Un lecteur fâché** de notre Tumblr épinglant les sites web de labo kitschs, réagissant sur le site *Educpros* :

Que c'est petit. Tout n'est pas que communication dans la vie. Qu'importe que le design soit démodé. L'important c'est qu'il fonctionne pour le public visé et, non, ce public n'est pas constitué de ceux qui s'en moquent. En ce qui concerne la page de Martin Karplus, donné en exemple, personnellement je préfère celle de Strasbourg, bien mieux centré sur lui, ses recherches, ses publications à celle de Harvard où on est noyé dans un fatras d'informations sur tout ce qui se fait là bas ... Se moquer -mal- d'individus pour vendre son propre business de com' ... pathétique

**Un autre lecteur** en désaccord avec le précédent réagit sur le site *Educpros* :

Belle initiative ! Rassurant qu'il y ait un poil d'irrévérente critique constructive dans l'air ! Non, je ne pense pas qu'il suffise que le design convienne à ceux qui ont fait la page. Sinon à faire une recherche qui n'intéresse et ne mobilise que les chercheurs ? On ne peut pas justifier l'indigence d'une page [web] par cet argument. Si ces sites [annuaires de chercheurs ou sites de laboratoires] sont si pauvres dans leur design, c'est que personne en France n'en a fait un enjeu national, une culture, comme c'est le cas aux Pays Bas où tout le monde est conscient des enjeux d'une bonne communication visuelle et où le bon designer est aussi respecté que le bon médecin. Le Design Graphique et de données servent précisément à ça, rendre clair et intelligible des contenus quels qu'ils soient de façon à ce qu'ils soient accessibles au plus grand nombre, compréhensibles et valorisés dans leurs singularités.

## BIBLIOGRAPHIE

### Rapports et articles académiques

- Académie des sciences (2012), *Remarques et propositions sur les structures de la recherche publique en France*, 25 septembre, [disponible en ligne](#)
- AMUE (2014), *Rapport d'activité 2013*, [disponible en ligne](#)
- Barré, Rémi, Yves Farge, Marie-Aline Bloch, Bernard David, Pascal Bain et Philippe Larrue (2007), Conduite du changement et qualité de la recherche publique, *Management de la recherche. Enjeux et perspectives*, pp. 157-242, [disponible en ligne](#)
- Bolton, Stuart (2010), *The Business Case for the Adoption of a UK Standard for Research Information Interchange. Report to JISC*, Stuart Bolton: Solutions, [disponible en ligne](#)
- Bordons, Maria et Isabel Gomez (2000), « Collaborations Networks in Science », dans *The Web of Knowledge: A Festschrift in Honor of Eugene Garfield*, pp. 197-214, dirigé par Blaise Cronin et Helen Barsky Atkins, [disponible en ligne](#)
- Clements, Anna, Scott Brander, Valerie McCutcheon, Josh Brown, Dale Heenan et Thomas Vestdam (2012), « CERIF in Action: Synthesise, standardise and productionise CERIF for Higher Education Institutions », *CRIS 2012: e-Infrastructures for Research and Innovation – Linking Information Systems to Improve Scientific Knowledge Production – 11th International Conf. on Current Research Information Systems*, Prague, 6-9 juin, pp. 279-288, [disponible en ligne](#)
- Dassa, Michèle et Isabelle Sidéra (2011), « Présentation du système d'information RIBAC du CNRS. Un outil de valorisation de l'activité des acteurs de la recherche en SHS », Colloque « Évaluation des productions scientifiques: des innovations en SHS ? », 9-10 juin, Paris, [disponible en ligne](#)
- Inspection générale de l'administration de l'Éducation nationale et de la Recherche (2013), *Audit de l'Agence de mutualisation des universités et des établissements*, décembre, [disponible en ligne](#)
- Katz, J. Sylvan et Ben R. Martin (1997), « What is research collaboration? », *Research Policy*, vol. 28, pp. 1-18, [disponible en ligne](#)
- Lee, Danielle H., Peter Brusilovsky et Titus Schleyer (2011), « Recommending collaborators using social features and MeSH terms », *Proceedings of the American Society for Information Science and Technology*, vol. 48, n° 1, pp. 1-10, [disponible en ligne](#)
- Lee, Danielle H. et Titus Schleyer (2010), « MeSH term explosion and author rank improve expert recommendations », *AMIA Annual Symposium Proceedings*, vol. 2010, pp. 412-416, [disponible en ligne](#)
- Ministère de l'économie et des finances (2014), *Rapport sur les politiques nationales de recherche et de formations supérieures. Annexe au projet de loi de finances pour 2015*, [disponible en ligne](#)
- Pumain, Denise et Frédéric Dardel (2014), *L'Évaluation de la recherche et de l'enseignement supérieur. Rapport à Madame la Ministre de l'enseignement supérieur et de la recherche*, janvier, [disponible en ligne](#)
- Russell, Rosemary (2012), *Adoption of CERIF in Higher Education Institutions in the UK: A Landscape Study*, UKOLN ISC, University of Bath, 15 mars, [disponible en ligne](#)
- Schleyer, Titus, Heiko Spallek, Brian S. Butler, Sushmita Subramanian, Daniel Weiss, M. Louisa Poythress, Phijarana Rattanathikum et Gregory Mueller (2008), « Requirements for expertise location systems in biomedical science and the Semantic Web », *International Semantic Web Conference 2008*, [disponible en ligne](#)
- Waddington, Simon, Allan Sudlow, Karen Walshe, Rosa Scoble, Lorna Mitchell, Richard Jones et Stephen Trowell (2013), « Feasibility Study Into the Reporting of Research Information at a National Level Within the UK Higher Education Sector », *New Review of Information Networking*, vol. 18, n° 2, [disponible en ligne](#)
- Wissenschaftsrat (2013), *Empfehlungen zu einem Kerndatensatz Forschung*, 25 janvier, [disponible en ligne](#)

## À PROPOS DES AUTEURS

### Antoine Blanchard

Après un stage de bibliométrie à l'[Observatoire des sciences et techniques](#) et un diplôme d'ingénieur agronome, Antoine a suivi le [Master « Etudes sociales des sciences et des technologies »](#) (Université Louis-Pasteur, Strasbourg) pour mieux comprendre et analyser les enjeux de la recherche contemporaine et de la science en train de se faire. Ses engagements en faveur du dialogue science-société l'ont d'abord rapproché des associations [Paris Montagne](#), Le Prisme à idées, [Traces...](#) jusqu'à ce qu'il en fasse son métier.

Après un début de carrière dans l'industrie, il a en effet créé [Deuxième labo](#) avec Elifsu pour façonner de nouvelles formes de relations science-société, et accompagner les mutations du monde de la recherche et de l'innovation. Aujourd'hui, en tant que chargé de programmes à l'université de Bordeaux, il s'occupe des actions de transfert de technologie et de valorisation des connaissances pour l'IdEx Bordeaux, à l'instar des Centres d'innovation sociétale. Il suit également les chantiers relatifs au numérique, dont le libre accès avec une [feuille de route remarquée](#).

Sur internet, il est actif depuis [plus de 15 ans](#) sous le pseudonyme d'Enro, et a été un pionnier des blogs de science dès 2006. Cette même année il a contribué à lancer la communauté de blogs du [C@fé des sciences](#), et en 2008 son association qu'il a présidée jusqu'en 2013. Depuis la rentrée 2014, il coordonne le Module « Culture scientifique et technique sur le web » dans le cadre du DU « Médiation scientifique innovante » à l'université Paris Diderot.

### Elifsu Sabuncu

Elifsu est docteure en biologie et intéressée par les questions scientifiques "chaudes" : maladies à prions (thèse de doctorat à l'Inra), usage des antibiotiques en France et mise en place de la vaccination contre le HPV (post-doc en épidémiologie à l'Institut Pasteur), perturbateurs endocriniens (post-doc en sociologie et histoire des sciences à l'Inra). Voir ses [publications...](#)

Également impliquée dans de nombreux enseignements, de la biologie aux humanités numériques (Master d'histoire à l'université du Luxembourg), en passant par les controverses scientifiques (Mines ParisTech et Sciences Po Paris), Elifsu est spécialiste de questions scientifiques à l'interface science-société. Elle a également participé à la naissance d'initiatives visant à rapprocher ces deux domaines : [festival de science Paris Montagne](#), programme [Science Académie](#) qui fait venir des lycéens au contact de la recherche, Club CNRS pour les jeunes.

Elifsu a co-fondé [Deuxième Labo](#) avec Antoine, pour contribuer à l'émergence de nouvelles formes de relations science-société et accompagner les mutations du monde de la recherche et de l'innovation. Dans ce cadre, elle a apporté son expérience de la participation des publics et des formats d'animation innovants pour faire vivre de nombreux projets collaboratifs. Elle est à l'origine, avec Antoine et le chercheur [Tom Roud](#), de l'initiative [Votons pour la science](#) qui a interrogé les candidats à la présidentielle française de 2012 sur leurs positions en matière de recherche.

Aujourd'hui, elle accompagne les associations Animath et Femmes et Mathématiques dans l'organisation des journées « Filles et maths : une équation lumineuse » et le Comité international des jeux mathématiques dans l'organisation du 16e Salon de la culture et des jeux mathématiques, place Saint-Sulpice à Paris.

À PROPOS DE DEUXIÈME LABO

# 2E LABO\_

**Pendant plus de 5 ans** et sous la direction d'Elifsu Sabuncu et Antoine Blanchard, **Deuxième labo** s'est développé comme agence de service et de conseil pour façonner de nouvelles formes de relations science-société. Ainsi, nous avons mobilisé nos compétences pluri-médias et pluri-disciplinaires pour composer des solutions pointues au service de projets de société. Nos clients ont représenté tous les secteurs de la société de la connaissance, des ministères (Ministère chargé de l'environnement) aux laboratoires et organismes de recherche (CNRS, Inserm, Inra, IRSN, Université de Strasbourg), en passant par les entreprises innovantes (Collectis, Syngenta), collectivités locales (Communauté d'agglomération du plateau de Saclay, Mairie d'Ivry-sur-Seine, Paris Métropole), associations et sociétés savantes (Science Ouverte, Paris Montagne, S[cube], Animath, Biophilia Educational Program, C'génial), et centres de culture scientifique et technique (Universcience, Cap sciences, Jardin des sciences).

**Aujourd'hui**, Deuxième labo est un *think and do tank* qui se penche sur les mutations du monde de la recherche et de l'innovation. Frappés par l'impuissance de l'institution scientifique à montrer sa richesse, à s'ouvrir à la pollinisation croisée et à renouveler son fonctionnement, nous sommes plus que jamais convaincus de l'urgence à préparer la recherche de demain, autour de trois axes fondateurs :

- pour un/e chercheur/e qui est entendu/e et fait la différence
- pour un labo ouvert et visible
- pour une recherche fertile et fertilisée

Si vous êtes chercheur ou ingénieur de recherche, dirigeant(e) d'organisme, chargé(e) de valorisation, responsable de systèmes d'information, animateur-trice de communautés de recherche... et que vous vous reconnaissez dans ces valeurs et ces objectifs, alors **contactez-nous** ! Sans plus attendre, nous pourrions nous attaquer ensemble à l'extension du domaine de la recherche.

## MENTIONS LÉGALES

L'illustration de couverture **représente géographiquement l'intensité des collaborations scientifiques dans le monde**, chaque arête symbolisant une co-publication — le tout à partir de la base des publications scientifiques Scopus. Elle est © Olivier H. Beauchesne – 2011.

Toutes les illustrations qui n'ont pas été produites par Deuxième labo sont © leur propriétaire respectif.

Le reste du document est placé sous licence Creative Commons BY.