



Données de la recherche : rôle des professionnels IST à l'Inra

Odile Hologne

► To cite this version:

Odile Hologne. Données de la recherche : rôle des professionnels IST à l'Inra. 43. congrès de l'ADBU, Sep 2013, Le Havre, France. 51 p. hal-02807047

HAL Id: hal-02807047

<https://hal.inrae.fr/hal-02807047>

Submitted on 6 Jun 2020

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.



Données de la recherche : Rôle des professionnels IST à l'Inra

Odile Hologne, directrice déléguée à l'IST



2007 : Début de l'histoire - séminaire FuturIST

Un contexte en (r)évolution

- La gestion de la production scientifique est stratégique pour les institutions
- Il appartient aux scientifiques de maîtriser la diffusion de leur production
- La production française représente 5% de la production mondiale des publications
- La valorisation des publications prend tout son sens dans le contexte international, et dans le respect des pratiques disciplinaires
- Et si on parlait services ? :
 - Enregistrement, préservation, réutilisabilité des données scientifiques
 - Accès à l'information (document en texte intégral)
 - Gestion des laboratoires
 - Politique scientifique, bibliométrie, infométrie
- Un cadre technologique omniprésent
- Une indissociabilité grandissante entre données primaires de recherche et publication scientifique
- Une importance grandissante à la qualification des objets numériques

Francis ANDRE 3 mai 2007 inra



Données primaires de recherche

- Très forte sous utilisation des ressources numériques primaires : surprotection ou simplement mauvaise habitude...
 - Le partage des données primaires est le garant d'une science plus efficace
- "Because they were denied access to data, 28% of geneticists reported that they had been unable to confirm published research,"
[Journal of the American Medical Association](#), 2005
- La gestion des données primaires, c'est (ce sera) aussi le rôle des documentalistes spécialisées

Francis ANDRE 3 mai 2007 inra



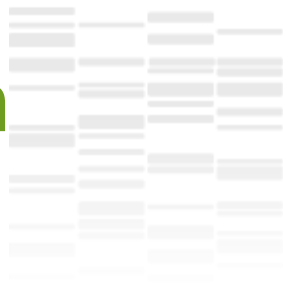
2013 : Aujourd'hui où en sommes nous ?

- ❖ L'Inra a une politique de partage des données
- ❖ Cette politique est mise en œuvre dans le cadre d'un projet qui implique une cinquantaine de personnes
- ❖ Les compétences mobilisées sont : scientifiques, informatique et IST
- ❖ L'IST est impliqué
 - ✓ Dans la mise en œuvre de la politique
 - ✓ Dans des projets avec les chercheurs
 - ✓ Dans des groupes de travail et des projets internationaux

Comment on en est arrivé là



Plan



- ❖ Un sujet politique et scientifique
- ❖ L'apport de l'IST
- ❖ Conclusion : compétences



01

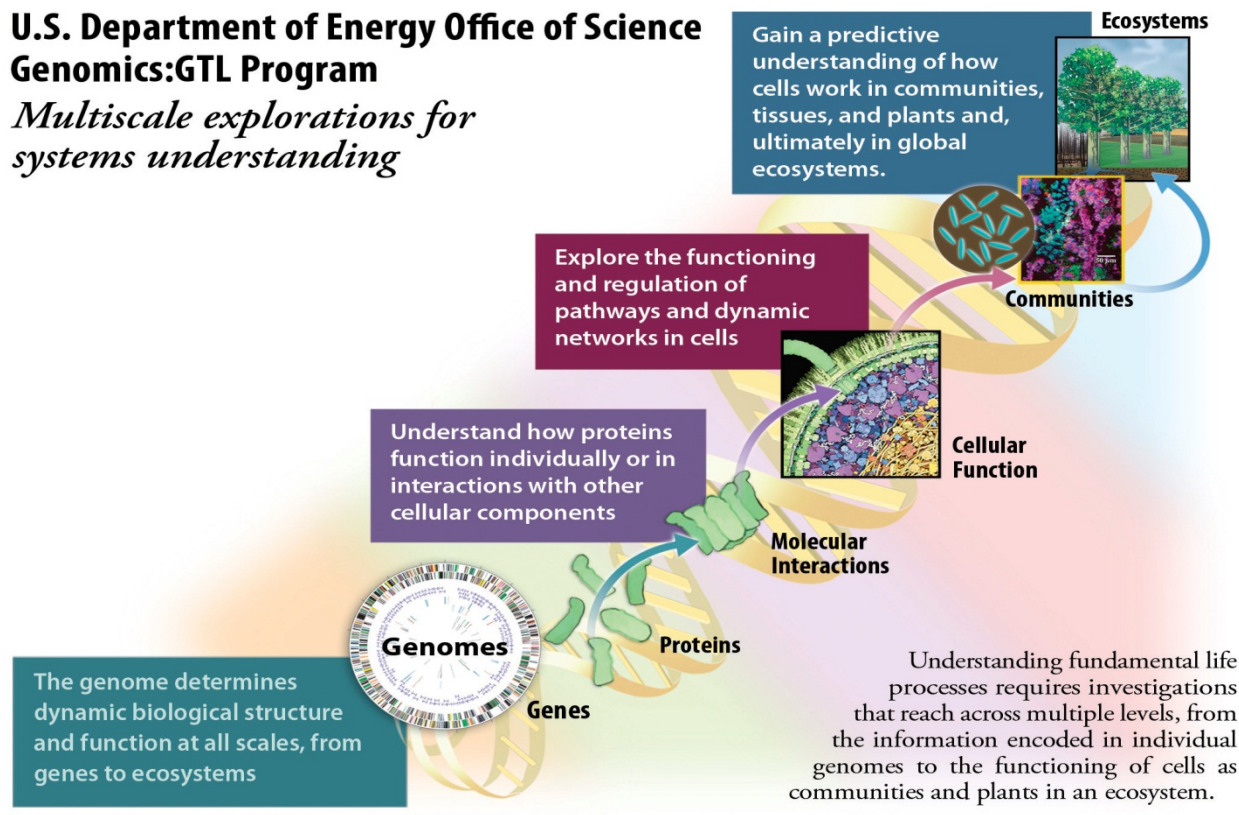
Les données un sujet politique et scientifique

Nous vivons une révolution technologique

❖ Une vision (presque) sans limite des possibles!

Rapprochement des thématiques dans un objectif de vision globale
=> Vers la biologie systémique

**U.S. Department of Energy Office of Science
Genomics:GTL Program**
*Multiscale explorations for
systems understanding*



Dans tous les domaines ...

Science, Février 2011 et

<http://research.microsoft.com/en-us/collaboration/fourthparadigm/>

IN SCIENCE

INTRODUCTION

Challenges and Opportunities

Science Staff

PERSPECTIVES

Climate Data Challenges in the 21st Century

J. T. Overpeck et al.

Challenges and Opportunities of Open Data in Ecology

O. J. Reichman

Changing the Visualization Ecology

P. Fox and J. H.

Challenges and Opportunities in Mining Neuroscience Data

H. Akil et al.

The Disappearing Third Dimension

T. Rowe and L. R. Frank

Advancing Global Health Research Through Digital Technology and Sharing Data

T. Lang

More Is Less: Signal Processing and the Data Deluge

R. G. Baraniuk

Ensuring the Data-Rich Future of the Social Sciences

G. King

Metaknowledge

J. A. Evans and J. G.

Access to Stem Cells and Data: Persons, Property Rights, and Scientific Progress

D. J. H. Mathews et al.

On the Future of Genomic Data

S. D. Kahn

NEWS FOCUS

Rescue of Old Data Offers Lesson for Particle Physics

A. Curry

Is There an Astronomer in the House?

S. Reed



Microsoft Research

Search Microsoft Research

Home Our Research Connections Careers Hub
About Us Research in Action Opportunities Research Accelerators

> Collaboration > The Fourth Paradigm: Data-Intensive Scientific Discovery > Table of Contents

Table of Contents

The Fourth Paradigm: Data-Intensive Scientific Discovery
Edited by Tony Hey, Stewart Tansley, and Kristin Tolle

This book is currently available as separate PDF files for online reading or downloading. Select chapters or individual essays below or get the [full-text version](#) of the book.

Introductions

- Foreword Gordon Bell
- Jim Gray on eScience: a transformed scientific method
Edited by Tony Hey, Stewart Tansley, and Kristin Tolle

Part 1: Earth and Environment

All Essays in Part 1

- Introduction Dan Fay
- Gray's laws: database-centric
Alexander S. Szalay, José A.
- The emerging science of environmental applications
Jeff Dozier, William B. Gail
- Redefining ecological science using data
James R. Hunt, Dennis D. Baldocchi, Catharine van Ingen
- A 2020 vision for ocean science
John R. Delaney, Roger S. Barga
- Bringing the night sky closer: discoveries in the data deluge
Alyssa A. Goodman, Curtis G. Wong
- Instrumenting the earth: next-generation sensor networks and environmental science
Michael Lehning, Nicholas Dawes, Mathias Bavay, Marc Parlange, Suman Nath, Feng Zhao

2009 : Présidente de l'Inra au BBSRC

Partage des données !

❖ Mais que fait l'Inra ?

➤ Question au conseil scientifique



BBSRC DATA SHARING POLICY

Version 1.1
June 2010

2011-2012: le CS instruit la question « Gestion et partage des données »



❖ Groupe de travail piloté par D. Pontier

❖ Objectifs

- ✓ Regard scientifique sur les spécificités thématiques en matière d'acquisition de données, de besoins en matière de gestion et de partage à l'Inra
- ✓ Regard sur les changements induits par les mouvements du Big Data et de l'Open Data
- ✓ Analyse de quelques cadres nationaux et internationaux existants

Gaspin, C., Pontier, D., Colinet, L., Dardel, F., Franc, A., Hologne, O., Le Gall, O., Maurin, N., Perrière, G., Pichot, C., Rodolphe, F. (2012). Rapport du groupe de travail sur la gestion et le partage des données <http://prodinra.inra.fr/record/206746>

2011-2012: le CS instruit la question « Gestion et partage des données »

❖ Conclusions (juin 2012): 9 recommandations

- ✓ R1 - Définir la politique de l'établissement et la communiquer
- ✓ R2 - Mettre en place un dispositif d'évaluation des données produites par l'Inra
- ✓ R3 - S'impliquer dans les comités internationaux de standardisation
- ✓ R4 - Développer un portail d'accès à un ensemble de ressources distribuées
- ✓ R5 - Prendre en compte le cycle de vie des données dès l'élaboration des projets de recherche

2011-2012: le CS instruit la question « Gestion et partage des données »

❖ Conclusions (juin 2012): 9 recommandations

- ✓ R6 - Définir un cahier des charges pour les plateformes
- ✓ R7 - Doter l'Inra d'infrastructures dimensionnées pour les stockages et calculs hautes performances
- ✓ R8 - S'engager dans une politique de gestion des compétences répondant aux besoins en émergence
- ✓ R9 - Conduire une réflexion inter-organismes

❖ Validation des recommandations par le CS en mai 2012

Comment bâtir notre politique « partage des données » ?

- ❖ Attendre ? NON
- ❖ Démarche « ingénieur » : réfléchir, groupe de travail ➡ aboutir à la politique
- Démarche stratégique : affirmer la politique
➡ mise en œuvre, groupes de travail
 - ✓ 3 rédacteurs : Informatique, scientifique, IST

Elaboration d'une politique

- ❖ 11 principes pour une **politique de partage des données de la recherche**
- ❖ Un projet inscrit au schéma directeur des systèmes d'information
 - ✓ création de data centers
 - ✓ Gestion prévisionnelles des emplois et compétences pour les sciences numériques

Avril 2013 : Lancement de la politique

- ❖ Séminaire avec : scientifiques, IST, Informatique
- ❖ Objectifs : Appropriation des principes et organisation du projet
- ❖ Résultats : création de groupes de travail :
 - ✓ Données par grands domaines
 - ✓ Méthodes et outils
 - ✓ IST
 - ✓ Stratégie : juridique, déontologie, communication
 - ✓ Compétences

État de l'art, étude d'opportunité dans le domaine IST

Contributeurs

Esther Dzale, Dominique L'Hostis, Pascal Aventurier, Véronique Garandel, Diane Le Hénaff, Virginie Lelièvre, + Sylvie Cocard

Périmètre

- Data journals : type d'articles, reviewing, stockage des données
- Mise en œuvre des DOI (Datacite Inist)
- Dépôt centralisé des données publiées : Prodinra2.0 ou autre alternative
- Citation des données : conseils aux auteurs
- Data citation index : Compréhension du fonctionnement, conseils à donner
- Repérage des bonnes ou mauvaises pratiques des éditeurs
- Évaluation des nvxx entrepôts de données externes (Figshare, Dryad, Pangaea)
- Annuaire des sources de données
- Certification des données

Principes de fonctionnement

Étude pour produire analyse des points du périmètre voire étude d'opportunité

Binômes à identifier en charge de chaque item
Échanges avec les axes données et méthodes-outils

Livrables - Jalons

- Chaque item fait l'objet d'un livrable. L'ensemble sera présenté au séminaire du 2 et 3 décembre 2013

Diapo n°47

Document de bilan du Séminaire de lancement avril 2013

Charte de déontologie

Les données, les savoir-faire et collections, la propriété industrielle : un patrimoine de qualité à préserver et valoriser

9. Les données de la recherche constituent l'une des composantes identifiables des productions de l'Institut, qu'elles soient issues des équipes de recherche, des infrastructures de recherche, des services administratifs. Elles sont la propriété de l'INRA, ou la copropriété, si elles sont obtenues dans le cadre d'accords de partenariat. La production, l'archivage, le traitement, la gestion des données de la recherche, obtenues par l'Institut ou transmises par des tiers, doivent obéir à des procédures explicitées, respectant les règles juridiques et déontologiques qui prévalent, notamment dans l'expérimentation ou dans l'utilisation des données personnelles. Ces procédures garantissent leur qualité, leur sécurisation, en particulier informatique, leur traçabilité, et fixent les règles de leur mise à disposition.

Charte de déontologie de l'INRA

Préambule

L'Institut national de la recherche agronomique (INRA) est un organisme public de recherche fondé, ouvert à l'international, conduisant des recherches, souvent en collaboration avec d'autres organismes partenaires ou avec des tiers. Ces recherches portent sur des enjeux sociétaux relevant de trois domaines principaux, l'agriculture, l'environnement et l'alimentation, interagissant avec les domaines liés que sont l'énergie, la chimie, la santé, les territoires. Les questions de recherche, complexes, pour lesquelles la société est en attente de réponses, exigent, à chaque instant, aussi bien de mettre en œuvre une démarche scientifique rigoureuse, que de s'interroger sur les enjeux éthiques qui en découlent. Pour ce faire, le comité consultatif d'éthique INRA-CERAD est en appui de la direction de l'Institut et de l'ensemble des personnels pour susciter le questionnement éthique sur les finalités des programmes de recherches et y apporter un éclairage.

L'INRA réalise les missions de service public qui lui sont confiées : la production, la diffusion et la valorisation de connaissances nouvelles, considérées comme biens publics, l'expertise, notamment en appui des politiques publiques, la contribution à la formation et à l'innovation, la communication vers la société. En tant qu'établissement public qui mobilise des collectifs de recherche, d'appui et de soutien à la recherche, l'INRA a la responsabilité de servir l'intérêt général. Pour ce faire, l'Institut privilégie la recherche de la pertinence, de l'utilité, de la qualité des résultats dans le respect mutuel, en mettant en avant l'impartialité et l'exemplarité de ses comportements, et en faisant preuve de vigilance sur ces questions.

C'est de la responsabilité de l'établissement que de donner des repères déontologiques à l'ensemble de ses personnels pour que chacun puisse inscrire, au cœur de ses missions et activités, l'ambition de servir cet intérêt général de façon impartiale et exemplaire.

La charte de déontologie exprime le sens général et les principes que chacun, personnellement, s'engage à respecter dans le cadre des missions qui lui incombent. Elle s'inscrit dans une démarche continue d'amélioration des pratiques professionnelles. Elle est portée par la direction générale de l'Institut et s'adresse à chaque agent qui y travaille, y compris à titre temporaire, l'établissement s'assurant que tous en ont pris connaissance. Elle est adossée à un ensemble de règles législatives, réglementaires ou institutionnelles qui régissent les activités professionnelles de tous. La loi n°83-634 modifiée du 13 juillet 1983 portant droits et obligations des fonctionnaires, le Code de la recherche, le Code de l'éducation, la Charte européenne de chercheur (11 mars 2005)¹, la Déclaration de Singapour sur l'intégrité en recherche (2010)² sont quelques-uns des principaux textes constituant le cadre qui promeut l'exercice de la responsabilité individuelle de chaque agent.

En faisant preuve, à l'échelle individuelle et collective, d'un comportement qui met en pratique la charte de déontologie dans toutes ses activités, l'Institut exprime aussi l'ambition de consolider le lien de confiance qui le lie à la société.

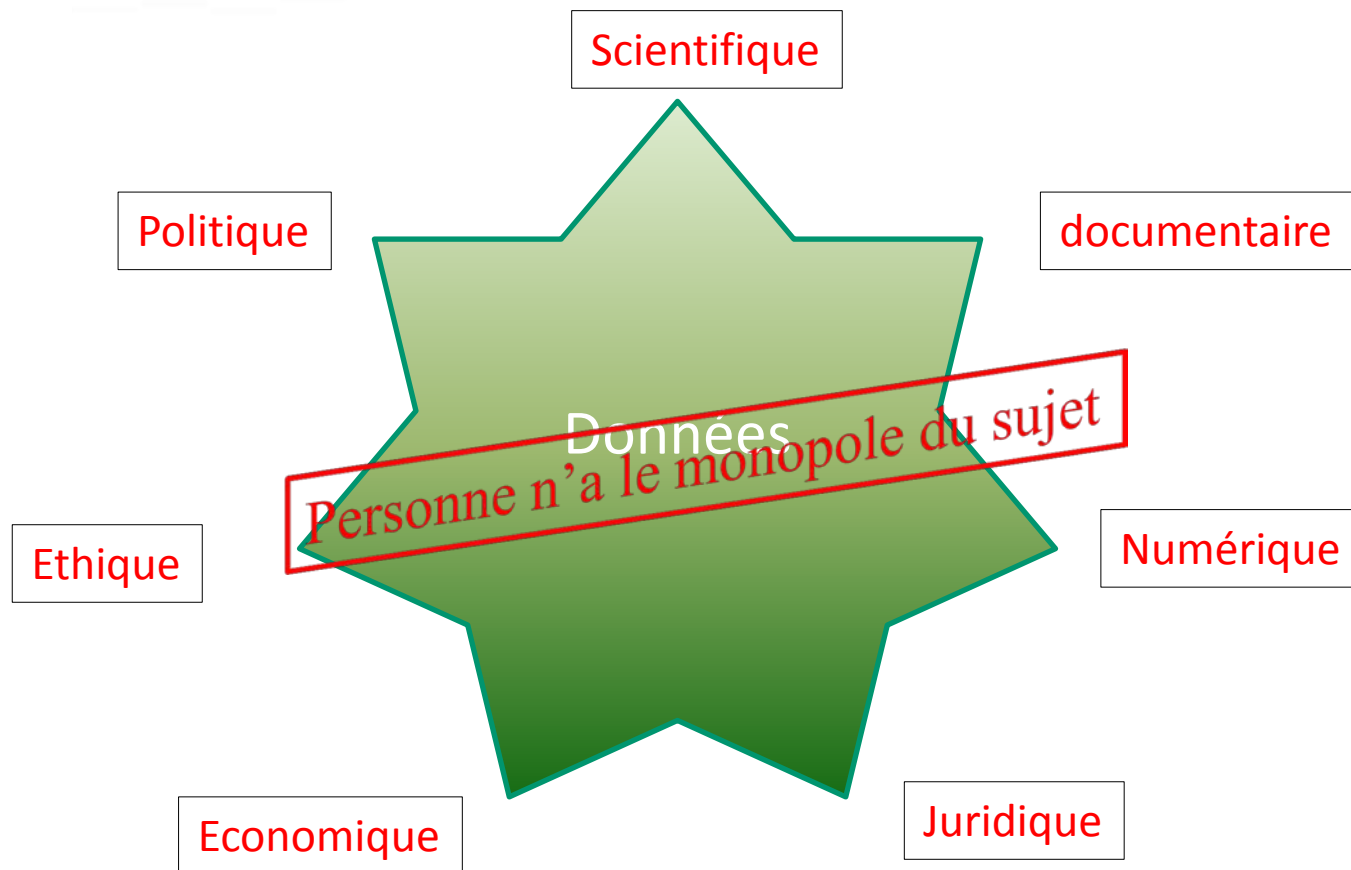
L'engagement de la direction générale de l'INRA

1. La direction générale de l'INRA définit et explicite les orientations, dans les champs scientifique et de l'appui à la recherche, qui seront mises en œuvre et les procédures qui seront suivies. Elle veille à prendre ses décisions avec impartialité et transparence et sera garante de l'application de ce principe à l'ensemble de l'Institut.

¹ http://ec.europa.eu/science/en/pdf/sci_21620_fr-fr.pdf

² http://www.singaporestatement.org/translation/05_French.pdf

Conclusion : un objet complexe



Des catalyseurs récents

❖ G8+5 Open data for Agriculture (avril 13)

Obtain commitment and action from nations and relevant stakeholders **to promote policies and invest in projects that open access to publicly funded global agriculturally relevant data streams**, making such data readily accessible to users in Africa and world-wide, and ultimately supporting a sustainable increase in food security in developed and developing countries



❖ Research data Alliance (mars 13)

The Research Data Alliance aims **to accelerate and facilitate research data sharing and exchange**





04

Apport des professionnels IST

Apport des professionnels IST

- ❖ Sur la thématique publication-données
- ❖ Sur l'appui à la définition de politique, dimension internationale
- ❖ Sur des aspects plus techniques (linked data)

Opportunités : Research Data Alliance

- ❖ Contribution à la création de 2 groupes (intérêt, travail)



The Research Data Alliance aims **to accelerate and facilitate research data sharing and exchange**

Agricultural Data Interoperability IG



Status: Recognised & Endorsed

The Agricultural Data Interest Group is a domain oriented interest group to work on all issues related to data important for the development of global agriculture. The interest group aims to represent all stakeholders producing, managing, aggregating, sharing and consuming data for agricultural research and innovation.

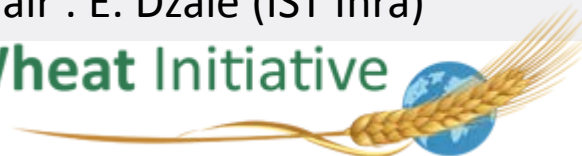
Wheat Data Interoperability WG



The Wheat Data Interoperability Working Group aims to provide a common framework for describing, representing linking and publishing Wheat data with respect to open standards.

Co chair : E. Dzalé (IST Inra)

Wheat Initiative



Autres opportunités internationales

- ❖ Participation au groupe « Research data » de Science Europe



- ❖ Organisation d'un meeting avec les partenaires de nos projets FP7 pour préparer H2020



Exemple d'intervention sur un projet technique : publication de données sur le web sémantique

- ❖ Rentabiliser l'investissement sur le web sémantique depuis 2 ans
- ❖ Montrer la plus value de ces technologies
- ❖ Montrer l'apport de l'IST
- ❖ Présentation de E. Dzalé au Forum GFII

http://forum.gfii.fr/uploads/docs/EstherKabor%C3%A9_websem_ForumduGFII.pdf

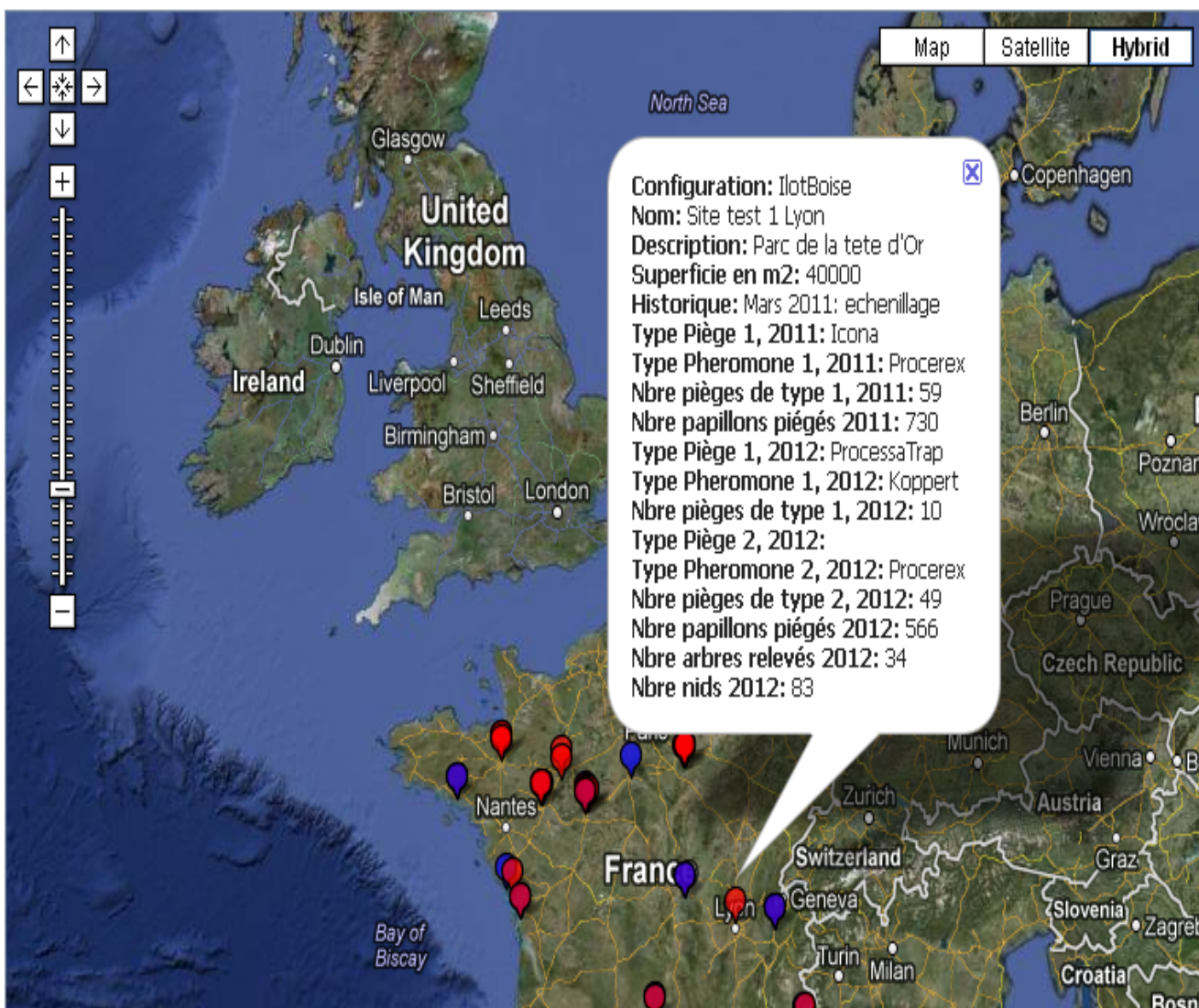
A	B	C	D	E	F	G	H	I	J	K	L
	Communes	Saint-Fargeau-Ponthierry	Obernai	Angers							Saint-Martin-de-Ré
	Nombre de pins et/ou de cèdres	96 pins	20 pins	97 pins 3 cèdres	18 pins	66 pins	22 pins	11 pins 3 cèdres	6 pins	31 pins	261 pins
	Hauteur des arbres	15 à 20 mètres	non relevé	10 à 15 mètres	10 à 25 mètres (avec un seul à 5 mètres)	non relevé	3 mètres à 20 mètres	5 mètres à 15 mètres	15 mètres à 20 mètres	4 mètres à 15 mètres	2 mètres à 10 mètres
Pièges à phéromones été 2011	Relevés des nids réalisés à l'hiver 2011	non réalisés	non réalisés	non réalisés							non réalisés et témoins
	Nombre de pièges installés	10	15	14 a) 1 b)	9	12 a) 1 b)	12 a) 2 b)	8 a) 1 b)	6 a) 1 b)	8	18
	Type de pièges	Procerex	Mastrap	a) Mastrap b) Nufarm	Mastrap	a) Mastrap b) Nufarm	a) Mastrap b) Nufarm	a) Mastrap b) Nufarm	a) Mastrap b) Nufarm	Mastrap	Mastrap
	Type de phéromones	Procerex	Procerex	Procerex	Procerex	Procerex	Procerex	Procerex	Procerex	Procerex	Procerex
	Date de la pose des pièges	25 juillet (sem 30)	01 juin (sem 22)	01 juillet (sem 26)			01 juillet (sem 26) et 06 juillet (sem 27)		31 mai (sem 22) et 19 juillet (sem 29)	19 juillet (sem 29)	12 juin (sem 2)
	Premiers vols de papillons constatés	non communiqué	Sem 26 et 27	sem 25							non relevé

Sites d'observation

69 SiteObservation

CARTE • DETAILS

6 résultats sur 69 ne peuvent pas être tracés.



Search

Localisation

- 2 Aix-les-Bains
- 8 Angers
- 2 Arzon
- 4 Avignon
- 3 Bayonne
- 2 Entraunes

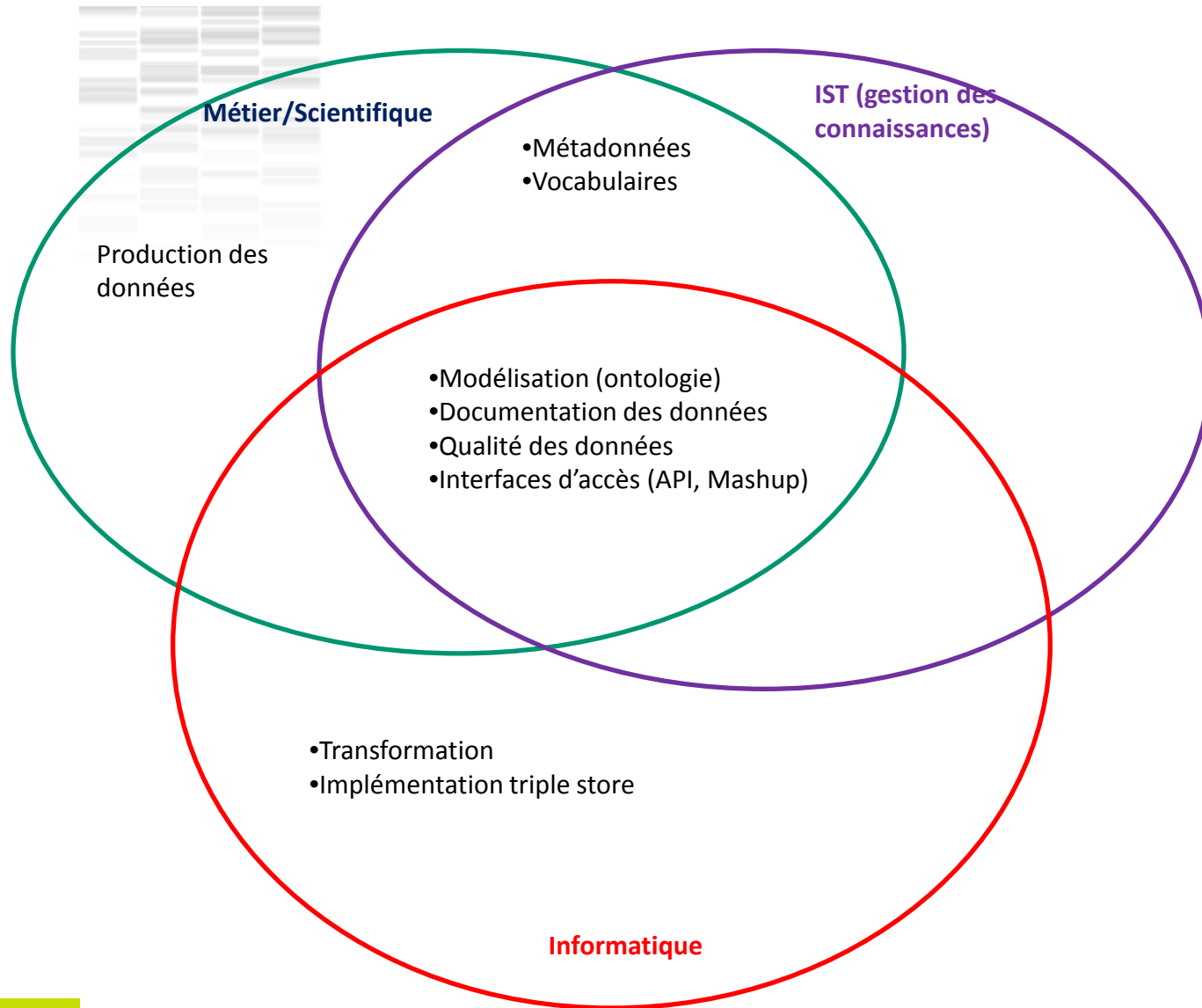
Configuration

- 1 (missing this field)
- 3
- 14 Alignement
- 3 ArbreSeul
- 48 IlotBoise

Type Site

- 23 Temoin
- 46 Test

Compétences mobilisées



Nouveaux challenges pour l'IST

NATURE | NEWS FEATURE

Publishing frontiers: The library reboot

As scientific publishing moves to embrace open data, libraries and researchers are trying to keep up.

Richard Monastersky

27 March 2013

The new data wranglers

.... At Johns Hopkins and many other top universities, libraries are aiming to become **more active partners in the research enterprise** — altering the way scientists conduct and publish their work. Libraries are looking to assist with all stages of research, by offering guidance and tools for collecting, exploring, visualizing, labelling and sharing data. “I see us moving up the food chain and being **co-contributors to the creation of new knowledge**,” says Sarah Thomas, the head of libraries

THE DATA GAP

A survey of more than 1,300 US scientists in 2010 showed an appetite for sharing data but significant hurdles that kept many from doing it.

67%

Said lack of access to others' data is a major impediment to scientific progress.

39%

Said their organization has an established process for long-term data storage.

22%

Said their organization or project provides funds for long-term data management.

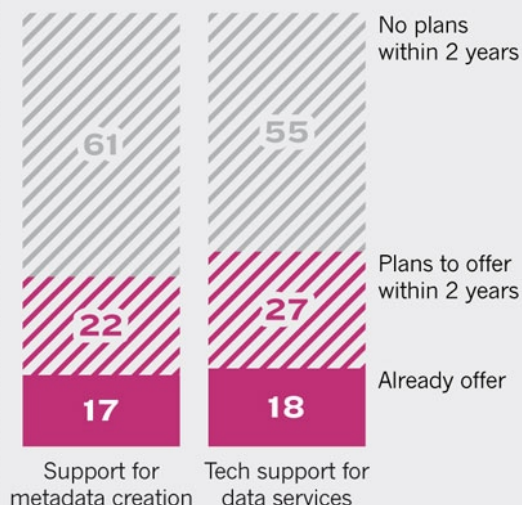
84%

Scientists who said that they would use other researchers' data sets if they were easily available.

36%

Scientists who said that others can access their data easily.

A survey of US research universities showed that many libraries are gearing up to provide data services.



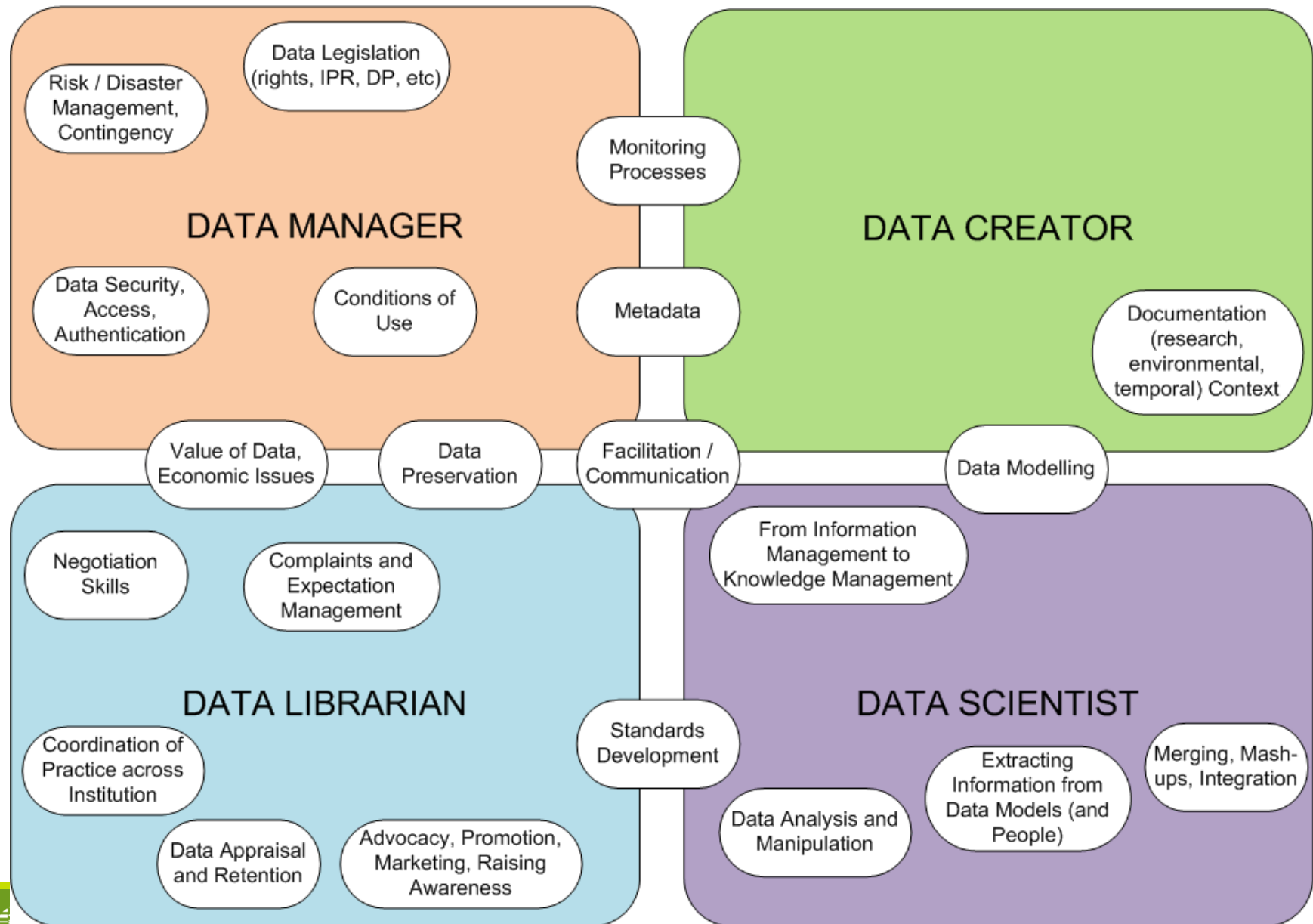


03

Compétences IST et données

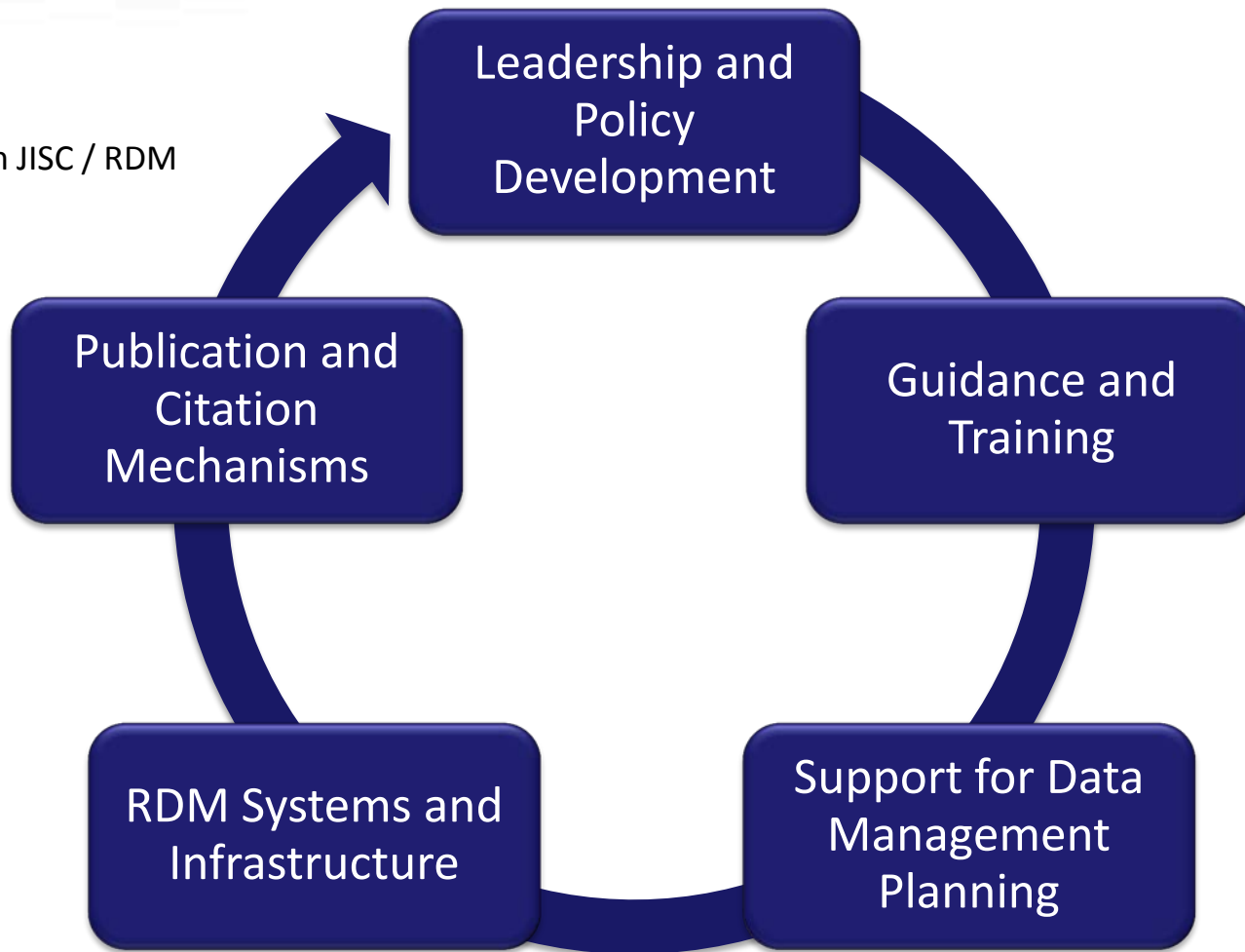
CORE SKILLS FOR DATA MANAGEMENT

A follow-up from the second DCC Research Data Management Forum (November 2008)



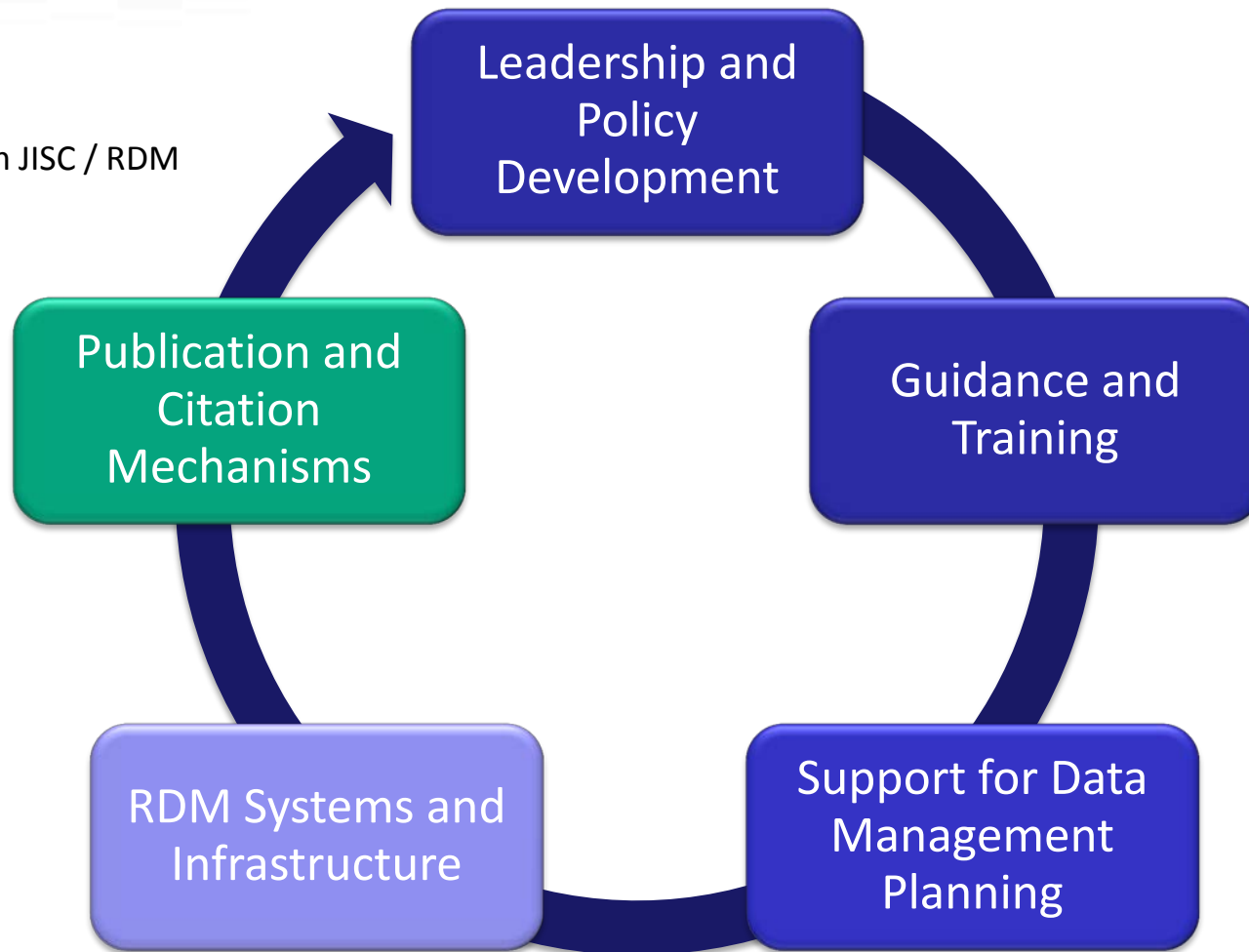
Gestion et partage des données : Différents domaines

Source : S. Hodson JISC / RDM
Data life cycle



Contribution possible de l'IST

Source : S. Hodson JISC / RDM



Compétences/connaissances à avoir

❖ Traditionnelles :

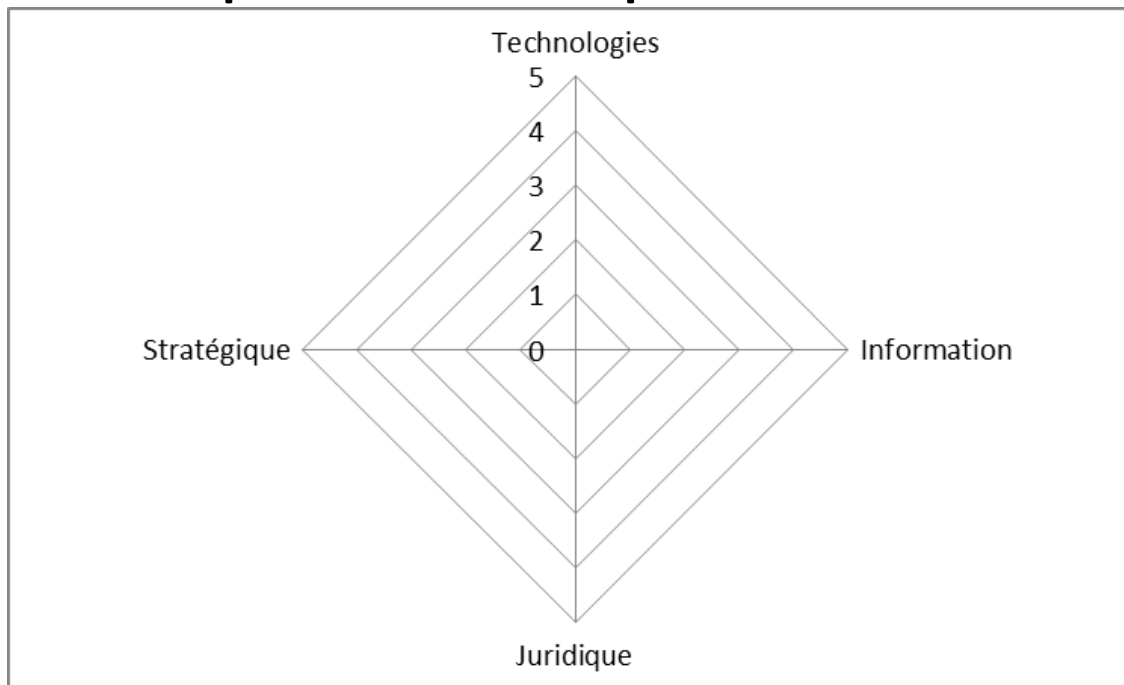
- ✓ compréhension des Data journals
- ✓ connaissance des entrepôts externes
- ✓ Mécanismes de citation et d'évaluation
- ✓ Connaissances juridiques (CC, PI, BdD)
- ✓ Maintenance de référentiels termino, vocabulaires ...

❖ Technologiques

- ✓ Identifiants (DOI, URI ...)
- ✓ Modèles de données et de métadonnées (DC et +)
- ✓ Interopérabilité (OAI-PMH et autres)
- ✓ Linked data, Ontologies et standards associés

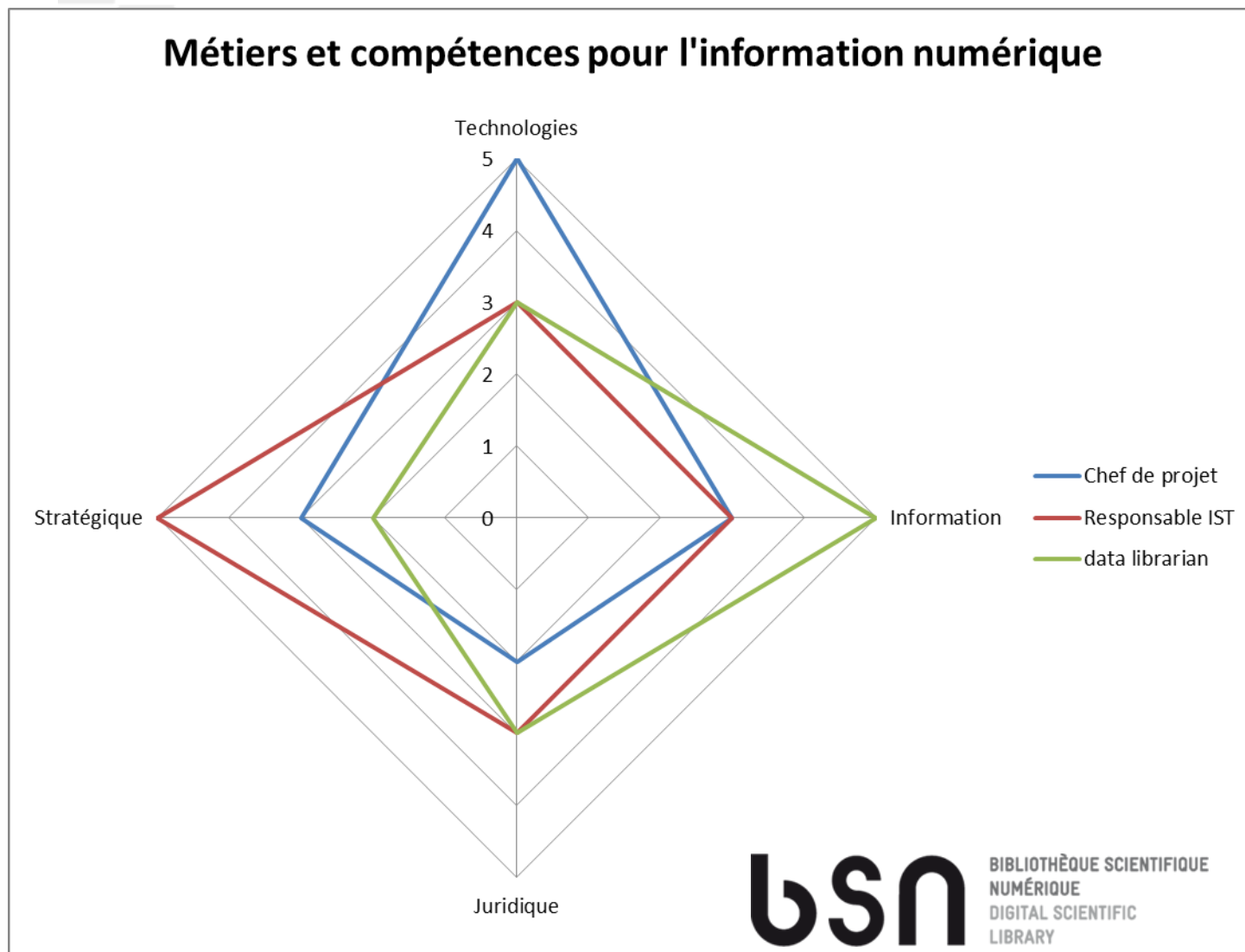
Travaux dans le cadre de BSN9

❖ 4 familles de compétences pour l'information scientifique numérique



bsn BIBLIOTHÈQUE SCIENTIFIQUE
NUMÉRIQUE
DIGITAL SCIENTIFIC
LIBRARY

Travaux dans le cadre de BSN9



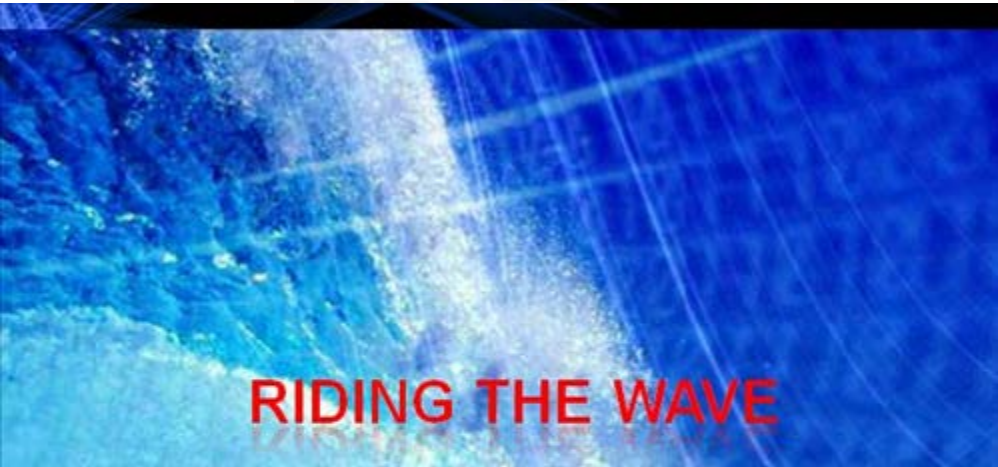
Pour conclure – A l'Inra, comment on en est arrivé là

- ❖ En montrant qu'on avait légitimité à parler
« données » avec notre cœur de métier (publi)
- ❖ En saisissant des opportunités, ou en les créant ...
- ❖ En ayant réinvesti depuis 2006 dans la fonction IST
 - ✓ Outils,
 - ✓ Formation,
 - ✓ Dispositif des pôles de compétences (dont 2 : gestion des connaissances et technologies IST)
 - ✓ Plan stratégique

Notre rôle en résumé

- ❖ Management : identifier les opportunités, trouver « une place » pour l'IST, bâtir des alliances (internes et externes), participation à l'élaboration de la politique « partage »
- ❖ Documentalistes : appréhender les nouveaux sujets, accompagner la politique auprès des communautés scientifiques
- ❖ Informatique/IST et IST/info : maîtrise technique LOD, démonstration des services
- ❖ Prochaines étapes :
 - ✓ Introduction du sujet « data » dans nos formations ?
 - ✓ IST et data centers -> data librarian ?
 - ✓ GPEC pour une science numérique

Surfer sur la vague



RIDING THE WAVE

**Les documentalistes ont
l'avenir devant eux,
mais... ils l'auront dans le dos
chaque fois qu'ils feront demi-tour !**

Laurent Bernat, Documentaliste SI, Volume 40, N° 2, paru le 30 avril 2003, page(s) 142-146

Lectures et sites intéressants



Researche Data Alliance <http://rd-alliance.org/>

Science as an open entreprise <http://royalsociety.org/policy/projects/science-public-enterprise/report/>

JISC <http://www.jisc.ac.uk/whatwedo/programmes/mrd.aspx>

Digital curation center

<http://www.dcc.ac.uk/>

<http://www.dcc.ac.uk/training/rdm-librarians> (formation pour les « Librarian »)

<http://www.dcc.ac.uk/events/idcc13/programme-presentations> Conf #idcc13

Australian national data service <http://www.ands.org.au/>



Merci pour votre attention...

odile.hologne@versailles.inra.fr

<http://prodinra.inra.fr/au/20625T>

@Holo_08  twitter





Annexes

Editeurs, Données, publications et évaluation

Ce que les professionnels IST doivent connaître
des évolutions en cours dans leur cœur de métier

En lien avec les publications

❖ Des nouvelles revues <http://proj.badc.rl.ac.uk/preparde/blog/DataJournalsList>



BGI Shenzhen and
BioMed Central

GigaScience aims to revolutionize data dissemination, organization, understanding, and use. An online open-access open-data journal, **we publish 'big-data' studies from the entire spectrum of life and biomedical sciences.** To achieve our goals, the journal has a novel publication format: one that **links standard manuscript publication with an extensive database that hosts all associated data and provides data analysis tools and cloud-computing resources.**

Not just 'omic' type data ... imaging, neuroscience, ecology, cohort data, systems biology and other new types of large-scale sharable data.

❖ Note aux auteurs – revues classiques



Data and materials availability All data necessary to understand, assess, and extend the conclusions of the manuscript must be available to any reader of *Science*.]...[

Science supports the efforts of databases that aggregate published data for the use of the scientific community. Therefore, appropriate data sets (including microarray data, protein or DNA sequences, atomic coordinates or electron microscopy maps for macromolecular structures, and climate data) must be deposited in an approved database, and an accession number or a specific access address must be included in the published paper. We encourage compliance with MIBBI guidelines (Minimum Information for Biological and Biomedical Investigations).

❖ De nouveaux entrepôts de données



Dryad is an international repository of data underlying peer-reviewed articles in the basic and applied biosciences. Dryad enables scientists to validate published findings, explore new analysis methodologies, repurpose data for research questions unanticipated by the original authors, and perform synthetic studies. Dryad is governed by a consortium of journals that collaboratively promote data archiving and ensure the sustainability of the repository.

JORD – programme du JISC

Journal Research Data Policy Bank (JoRD)

The Journal Research Data Policy Bank (JoRD) project will be conducting a feasibility study into the scope and shape of a sustainable service to collate and summarise journal policies on Research Data. The aim of this service will be to provide researchers, managers of research data and other stakeholders with an easy source of reference to understand and comply with Research Data policies.

Through maintaining a firm focus upon research literature and stakeholder consultations, this project will identify and consult with a wide range of stakeholders (nationally and internationally) and look at journal policies on Research Data from leading journals to deliver detailed requirements and specifications for a service that will provide a ready reference source of easily accessible, standardised, accurate and clear information relating to the journal policy landscape for research data. A major part of the work done by JoRD will focus upon researching and analysing a range of business and sustainability models and delivering a business framework that will allow any proposed service to continue on a long term sustainable basis.

The JoRD project will work with a wide range of partners including Research Information Network (RIN) and will draw upon advice and support from the DCC, Dryad, Australian National Data Service the International STM Publishers Association and a large network of international SHERPA RoMEO Partners and collaborators.

Summary

Start date

9 July 2012

End date

31 December 2012

Funding programme

Digital infrastructure:
Research management
programme

Strand

Innovative Research Data
Publication

Project website

Lead institutions

University of Nottingham 

Partner institutions

Research Information Network
(RIN) 

Respectful re-use

Nature Genetics **44**, 1073 (2012) | doi:10.1038/ng.2433

Published online 26 September 2012

The impact of the papers we publish depends increasingly on the data they describe. In insisting on data access for referees and readers, we prioritize scientific integrity above all and place the interests of research participants before impact.

The Nature research journals insist that materials, methods and data be made available and that authors detail any conditions for access where these exist. In our shared guide to authors (<http://www.nature.com/authors/policies/availability.html>), we state:

The preferred way to share large data sets is via public repositories. Some of these repositories offer authors the option to host data associated with a manuscript confidentially, and provide anonymous

[print](#)[email](#)[pdf](#)[download citation](#)[order reprints](#)[rights and permissions](#)

Helping you publish, discover,
and reuse research data



Welcome to *Scientific Data*

Scientific Data is a new open-access, online-only publication for descriptions of scientifically valuable datasets. It introduces a new type of content called the Data Descriptor, which will combine traditional narrative content with curated, structured descriptions of research data, including detailed methods and technical analyses supporting data quality. *Scientific Data* will initially focus on the life, biomedical and environmental science communities, but will be open to content from a wide range of scientific disciplines. Publications will be complementary to both traditional research journals and data repositories, and will be designed to foster data sharing and reuse, and ultimately to accelerate scientific discovery.

Scientific Data will launch in Spring 2014 and be open for submissions in Autumn 2013. Sign up for our

Lien avec les citations des articles

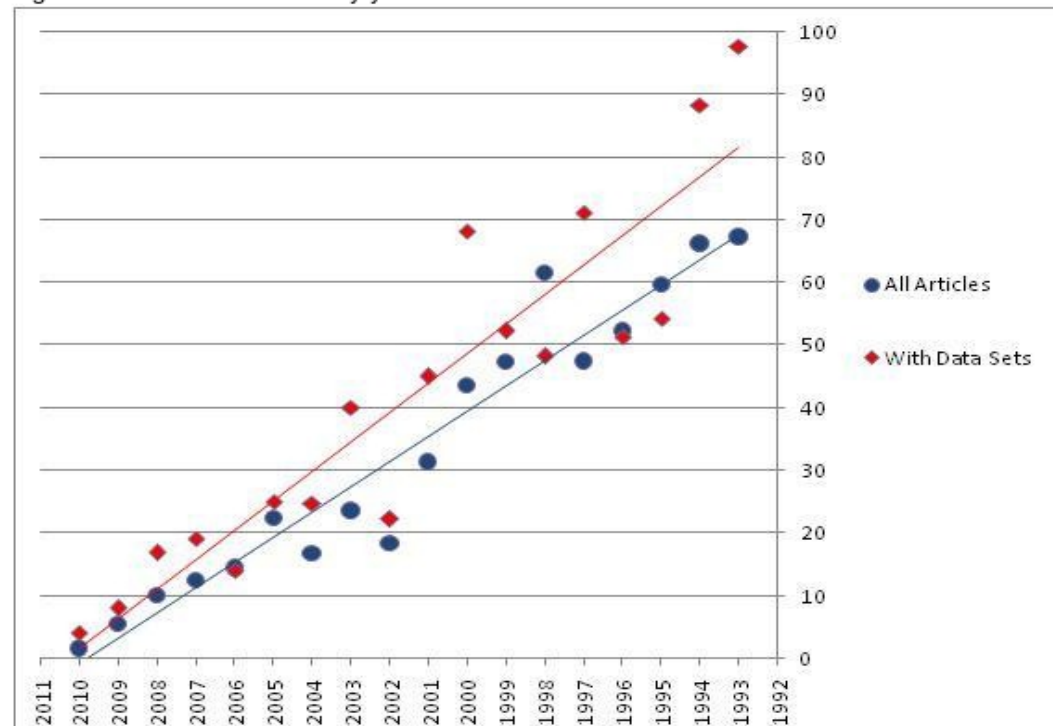
Data Sharing Effect on Article Citation Rate in Paleoceanography

posted Nov 27, 2011 4:11 AM by Michael Diepenbroek [updated Nov 27, 2011 11:21 PM]

Jon Sears (AGU) - Abstract for the AGU 2011:

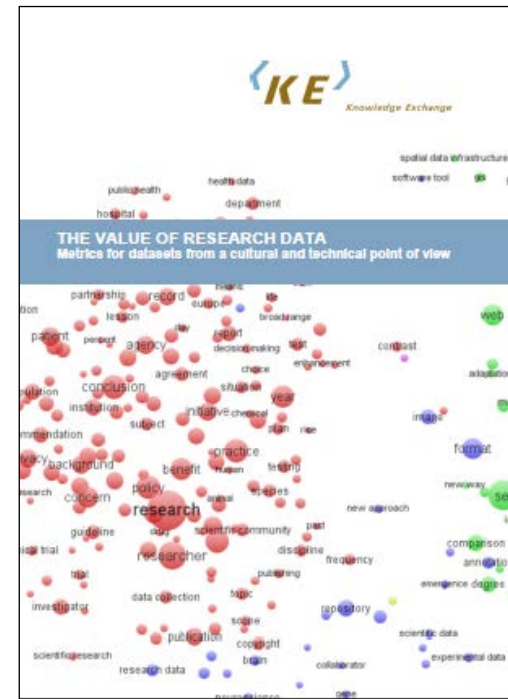
The validation of scientific results requires reproducible methods and data. Often, however, data sets supporting PANGAEA® data library measurably increases the citation rate of articles published between 1993 and 2010 in paleoceanography. Articles with available supporting data sets received 19.94% (8,056) of the aggregate citations (40,409). Publicly available citations were sampled over the 18-year study period (1,331), and the increase is fairly consistent over time (14 of 18 years).

Figure: Plot of mean citations by year to articles with data sets vs. citations to all articles in Paleoceanography between 1993 and 2010.



uer les

<http://www.knowledge-exchange.info/datametrics>



Codata (Committee on Data for Science and Technology): siège à Paris

Data Science Journal, Volume 12, 13 September 2013



Executive director
Simon Hodson
Twitter @simonhodson99

OUT OF CITE, OUT OF MIND:

THE CURRENT STATE OF PRACTICE, POLICY, AND TECHNOLOGY FOR THE CITATION OF DATA

CODATA-ICSTI Task Group on Data Citation Standards and Practices

Edited by Yvonne M. Socha

Building a Culture of Data Citation



THE DATA CITATION INDEXSM

DEFINITIONS:

Data repository: a database or collection comprising data studies, data sets and/or microcitations which stores and provides access to the raw data. Constituent data studies, and sometimes individual data sets, are marked up with metadata providing a context for the available raw data.

Data study: description of studies or experiments held in repositories with the associated data which have been used in the data study. (Includes serial or longitudinal studies over time). Data studies can be a citable object in the literature and may have cited references attached in their metadata, together with information on such aspects as the principal investigators, funding information, subject terms, geographic coverage etc. The level of metadata provided varies between repositories.

Data set: a single or coherent set of data or a data file provided by the repository, as part of a collection, data study or experiment. Data sets may present in a number of file formats and media types: they may be number based files such as spreadsheets, images, video, audio, databases etc. Data sets can be a citable object in the literature and may have cited references attached in their metadata, but more commonly they inherit the metadata of the overall study in which they are used.

- ☐ 1. Title: **ESTerases and alpha/beta Hydrolase Enzymes and Relatives.**
Editor(s): Hotelier, Thierry; Renault, Ludovic; Cousin, Xavier; et al.
Source: ESTerases and alpha/beta Hydrolase Enzymes and Relatives
Source URL: <http://bioweb.ensam.inra.fr/ESTHER/general?what=index>
Document Type: **Repository** Times Cited: **2** (from All Databases)
[ [View abstract](#)]



- ☐ 1. Title: **Enzymatic Activity and Protein Interactions in Alpha/Beta Hydrolase Fold Proteins: Moonlighting Versus Promiscuity**
Author(s): Marchot, Pascale; Chatonnet, Arnaud
Source: PROTEIN AND PEPTIDE LETTERS Volume: **19** Issue: **2** Pages: **132-143** Published: **FEB 2012**
Times Cited: **3** (from All Databases)
 [ [View abstract](#)]

- ☐ 2. Title: **ESTHER, the database of the alpha/beta-hydrolase fold superfamily of proteins**
Author(s): Hotelier, T; Renault, L; Cousin, X; et al.
Source: NUCLEIC ACIDS RESEARCH Volume: **32** Special Issue: **SI** Pages: **D145-D147** DOI: **10.1093/nar/gkh141** Published: **JAN 1 2004**
Times Cited: **79** (from All Databases)
  [ [View abstract](#)]

Figshare – Digital Science (MacMillan/Nature)



My data  Browse Upload

 O. Hologne ▼

My data

0% of private storage used

search titles, tags, authors, description, etc... 

☐ Add to Fileset Batch edit

	Type mouseover	Date	Status	Statistics public items only
☐ esi18012012.xlsx	 DATASET	25.01.2012 13:31	DRAFT	Add info

showing all 1 items