

# Les chercheurs peuvent-ils continuer à ignorer le droit ?

**Patrick Séchet**

L'activité de recherche scientifique proprement dite comprend souvent une phase d'expérimentation qui donne lieu à la collecte et au traitement d'une série de données d'observation. Vient ensuite une étape plus conceptuelle d'interprétation des faits observés qui suppose des études et comparaisons et, partant, l'exploitation d'un corpus élargi d'informations. Ainsi, le chercheur se trouve-t-il dans son activité principale tour à tour producteur et utilisateur de données scientifiques.

En outre, même si le rôle essentiel des équipes de recherche est de produire de la connaissance certifiée, c'est-à-dire validée par les pairs par le biais de la sélection d'articles à publier dans des revues scientifiques internationales, il leur est aussi de plus en plus demandé de contribuer à la diffusion de ces connaissances. Cela se traduit par diverses participations à l'innovation technologique, à la formation, à l'expertise, aux débats de société, etc., et de ce fait les chercheurs se trouvent plus qu'autrefois confrontés aux problèmes de la société.

## Propriété des données scientifiques

Pour toutes les activités mentionnées, la question de la propriété des données scientifiques et des connaissances produites se pose immédiatement. Elle est d'autant plus importante que l'évolution récente de notre société contribue largement à l'émergence d'un véritable marché de l'information, alors que parallèlement la connaissance devient un facteur crucial dans la compétition économique que se livrent les entreprises.

Les questions que l'on peut alors légitimement se poser sont variées. De quelles prérogatives l'équipe de recherche dispose-t-elle sur les données qu'elle produit dans le cadre de son activité ? Comment devenir utilisateur légitime pour les données scientifiques que l'on envisage d'exploiter ? Quelles utilisations sont alors possibles, quelles restrictions existe-t-il et quelles sont les obligations à respecter ? Ces problèmes sont d'autant plus complexes que l'environnement du chercheur fait généralement intervenir de multiples acteurs : l'organisme dont il relève, les autres membres de l'équipe, le partenaire institutionnel éventuel (souvent étranger dans le cas de l'IRD), les intermédiaires, les bailleurs de fonds, parfois des contractants, etc.

La première intervention, proposée par Sébastien Lafargue de l'Ifremer, livre un tour d'horizon des régimes juridiques applicables à la protection des données, dans le but de définir les droits et devoirs respectifs du créateur et de l'utilisateur. Entre autres choses, elle met en avant la distinction à faire entre données brutes et données élaborées.

## Diffusion des données scientifiques

Les informations qui ont été recueillies, parfois à grands frais, dans le cadre d'une opération de recherche déterminée, sont toujours utilisables dans un autre contexte et pour d'autres fins. Pour favoriser cette réutilisation elles doivent être mises à la disposition de la communauté scientifique, voire de la société tout entière, par le biais des moyens appropriés dont le plus commode actuellement est la base de données accessible en ligne.

Là encore, la simple question de la diffusion des données publiques n'est pas triviale sur le plan du droit, même si en France la circulaire du 14 avril 1994 <sup>1</sup> relative à la diffusion des données publiques donne le cadre général pour un organisme tel que l'IRD.

L'exposé préparé par Bernard Marx, de l'Inpi, fournit un exemple de mise en œuvre avec les informations sur les entreprises et les données de propriété industrielle dont cet organisme est le dépositaire et pour lesquels il a développé des services d'information appropriés, conformément à ses missions.

## Cas particulier des données nominatives

La nature même des données scientifiques peut parfois introduire d'autres contraintes d'ordre juridique. C'est par exemple le cas - souvent rencontré dans le domaine de la santé - des données nominatives. On sait en effet que la loi du 6 janvier 1978, dite informatique et libertés, avait imposé un certain nombre de contraintes pour la création de fichiers sur ce genre de données, afin de prévenir toute utilisation sensible au regard du respect du droit des personnes et de la vie privée.

L'intervention de Jeanne Bossi, de la CNIL, s'attache plus particulièrement à présenter les principes et les modalités applicables dans le cas des traitements de ce type de données induits par la recherche dans le domaine de la santé, notamment l'épidémiologie. Ceux-ci ont été explicités par une loi du 1<sup>er</sup> juillet 1994, venant compléter le régime général établi par la loi susmentionnée.

## Question de la fraude sur les données scientifiques

Enfin, même si le plagiat et les résultats tronqués ou inventés ont toujours existé, le sujet de la lutte contre la fraude sur les données scientifiques commence seulement à émerger de manière plus formelle. Dans la mesure où l'avancée des connaissances est un phénomène cumulatif, il est clair en effet que l'aspect confiabilité des résultats est essentiel et qu'il faut donc se donner les moyens d'éviter toute fausse piste.

Françoise Souyri, de l'Inserm, présente dans son intervention l'état actuel des connaissances dans ce domaine. Elle montre qu'il y a émergence d'une prise de conscience et de la nécessité de se plier à certaines règles (comité d'éthique, bonnes pratiques de laboratoire, etc.). Quelques pays ont déjà pris des dispositions au niveau organisationnel et, en France, cet organisme a récemment pris des mesures dans le même sens.

## Conclusion

Cette session n'avait évidemment pas la prétention d'être exhaustive : on s'est seulement efforcé d'aborder les questions juridiques soulevées à propos des données scientifiques. On constate alors que tous ces problèmes, qui apparaissent pourtant fondamentaux pour la recherche scientifique, ont longtemps été éludés et restent actuellement largement méconnus. Ils sont encore très souvent à l'origine de conflits et de déboires aujourd'hui.

Le fait que la communauté scientifique soit très mal informée sur le plan du droit semble la cause essentielle de ces problèmes. Qui par exemple, dans les organismes publics de recherche aussi bien que dans les universités connaît les lois de juillet 1983 et janvier 1984, propres à la fonction publique d'État ? Ce sont pourtant ces textes qui définissent les droits et obligations des fonctionnaires, et donc des personnels de la recherche publique. On y trouve en particulier les principes d'obéissance à la hiérarchie ou de l'obligation de réserve.

Bien d'autres textes législatifs et réglementaires devraient être connus des chercheurs, ne serait-ce que parce qu'ils y sont confrontés dans leur activité quotidienne. Ceux qui sont cités le plus souvent, comme la loi Informatique et libertés déjà évoquée, la loi du 4 août 1994 relative à l'emploi de la langue française (dite loi Toubon) ou la loi du 5 janvier 1988 relative à la fraude informatique (dite loi Godfrain) ne sont pas nécessairement les plus importants. En particulier tous les aspects liés à la notion de la responsabilité, notamment lorsque le chercheur est sollicité pour donner un avis d'expert, devraient être mieux connus.

D'ailleurs, tous les sujets qui sont ainsi abordés au cours de cette session montrent bien qu'au fond le juridique est omniprésent dans les activités du chercheur, quand bien même le droit est généralement ignoré par les intéressés. Les solutions qui sont proposées pour pallier cette ignorance consistent le plus souvent à élaborer des textes généraux, sorte de chartes auxquelles les équipes sont invitées à adhérer et à se conformer. Par exemple, s'agissant des données scientifiques, on trouvera en annexe l'esquisse d'une proposition de politique de l'institution<sup>2</sup>, censé régler la question générale de la propriété et de la diffusion des informations, dans le respect de la circulaire Balladur.

## Annexe

### Proposition de politique de l'IRD en matière de données scientifiques

La nécessité pour l'Institut de se doter d'une politique claire quant au devenir des données qu'il produit est évoquée de manière récurrente depuis plusieurs années. L'élaboration de cette politique s'est toutefois heurtée à des difficultés tenant à la fois à une connaissance insuffisante du cadre juridique (lui-même encore instable) et à la diversité particulièrement grande à l'IRD des situations rencontrées. Fort d'une meilleure assimilation des textes réglementaires en vigueur, au moins au niveau national et communautaire, on peut maintenant fournir les quelques éléments de discussion indispensables pour esquisser une telle politique. Sont préalablement exclues du champ de cette politique les informations qui doivent être maintenues au secret, par exemple lorsqu'il y a possibilité d'obtention d'un brevet.

#### Considérations préliminaires

1. Considérant que dans le strict cadre de ses programmes de recherche l'IRD collecte, traite et interprète un corpus important et varié de données scientifiques originales. De telles informations peuvent avoir d'autres utilisations que celle pour laquelle elles ont été initialement recueillies, de sorte que l'établissement a indéniablement un rôle à jouer dans la diffusion de ces connaissances.
2. Considérant toutefois que la mission de recherche scientifique pour le développement en coopération qui est assignée à l'Institut ne prévoit pas explicitement de mission spécifique de capitalisation et mise à disposition de données. De telles actions sont cependant ponctuellement justifiées pour que l'organisme puisse valoriser ses acquis dans le prolongement de sa mission.
3. Considérant que l'IRD, comme établissement public à caractère scientifique et technologique, relève du secteur public français. À ce titre et *a priori*, l'organisme n'est pas exclu du champ d'application de la circulaire du Premier ministre en date du 14 février 1994 sur la diffusion des données publiques.
4. Considérant qu'en l'état actuel de la doctrine internationale, un régime de droit privatif dit du « droit d'auteur » s'applique indifféremment aux données scientifiques élaborées et, plus généralement, à l'ensemble des produits informationnels associés (bases de données, logiciels, produits multimédias, etc.).
5. Considérant que du fait de la modalité de partenariat qui s'applique à la plus grande partie de ses activités, l'Institut agit le plus souvent dans un cadre contractuel. Les données sont donc collectées, et les informations obtenues, en régime de partage avec le(s) partenaire(s) institutionnel(s).
6. Considérant enfin que pour se ménager l'accès à des données scientifiques recueillies par d'autres organismes, l'IRD doit pouvoir offrir ses propres informations comme monnaie d'échange.

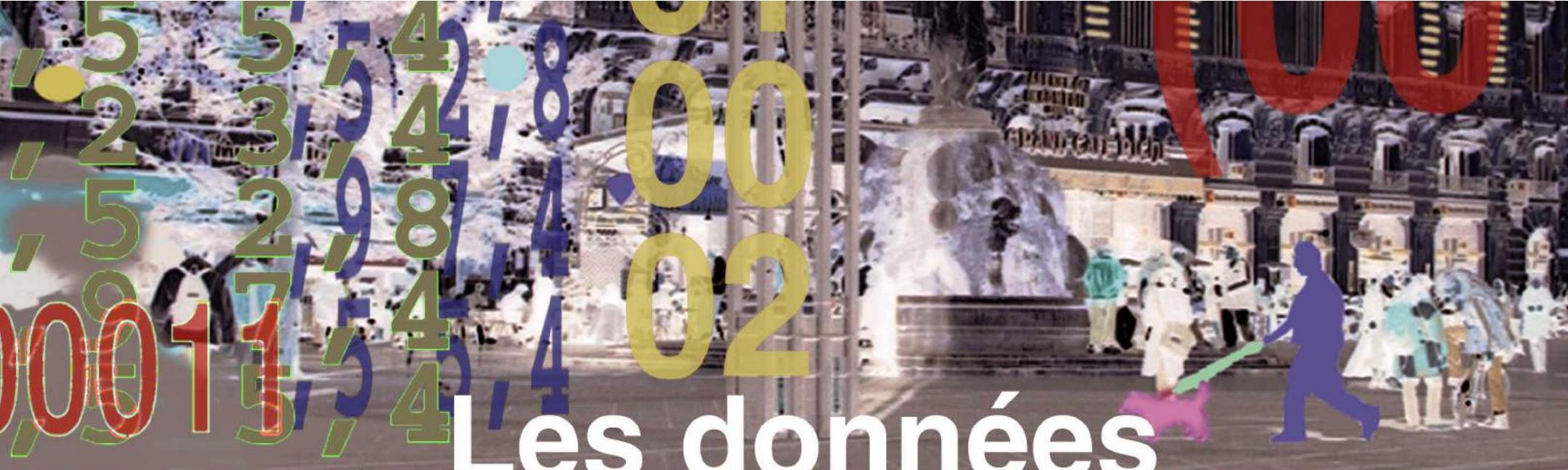
**L'IRD se fixe la politique suivante par rapport aux produits informationnels**

- a) les produits informationnels, et en particulier les données scientifiques collectées et traitées par les agents de l'Institut dans le cadre des missions qui leur sont dévolues, appartiennent à l'Institut, sauf disposition contractuelle explicite contraire. Cette règle est applicable aux personnels mis à disposition, sauf convention expresse. L'IRD s'engage à faire respecter les droits moraux correspondants de ses agents, et notamment la mention de paternité sur ces données ;
- b) l'Institut prévoit systématiquement dans ses programmes de recherche une étape devant conduire à la capitalisation et à la mise à disposition des produits informationnels générés par ces programmes. De règle, cette opération doit être externalisée (l'IRD pouvant y contribuer financièrement). Trois exceptions sont néanmoins tolérées : lorsque cette opération vise simultanément l'appropriation en interne d'un savoir-faire nouveau, lorsqu'il n'existe pas d'autres possibilités de capitaliser ces données ou lorsque celles-ci présentent une valeur stratégique pour l'IRD <sup>3</sup>.
- c) en aucun cas l'Institut ne cède ses droits patrimoniaux sur ses produits informationnels. La capitalisation et la mise à disposition des données de l'IRD par un tiers se font sur la base d'un contrat de concession d'un droit d'utilisation et de diffusion avec ce tiers. Cet accord doit prévoir systématiquement que l'Institut est utilisateur légitime pour la totalité du corpus de données auquel celles apportées par l'IRD ont été intégrées. Lorsque la diffusion est réalisée à titre onéreux, un pourcentage de redevances doit être prévu, l'IRD faisant son affaire d'éventuels reversements de « *royalties* » à ceux de ses partenaires qui ont participé à l'obtention de ces résultats ;
- d) pour assurer la visibilité de l'Institut et celle de ses partenaires, une mention d'origine doit accompagner toute extraction de ces données lorsqu'elles sont fournies par le diffuseur à un demandeur. Une telle mention doit être prédéfinie dans l'instrument contractuel prévu à l'alinéa précédent. En termes d'utilisation à des fins scientifiques ou d'illustration de l'enseignement, seules les exceptions explicitement prévues par la législation sur la protection juridique des bases de données sont admises *a priori*.

<sup>1</sup> Dite circulaire Balladur.

<sup>2</sup> Document établi en février 1996.

<sup>3</sup> Correspond par exemple à la notion de « collection de référence » d'une discipline.



# Les données scientifiques

Bases de progrès des connaissances  
*Séminaire tenu à l'IRD Ile-de France, Bondy les 4 et 5 mai 1999*

Éditeurs scientifiques Jean-Michel Kornprobst, Marcel Raffy

## Sommaire

### Introduction

Marcel Raffy, professeur, université de Strasbourg

Jean-Michel Kornprobst, professeur, université de Nantes

Jean-Pierre Muller, directeur général de l'IRD

### 1<sup>re</sup> partie : acquisition et stockage des données

*Coordinateur* : Francis Laloë, IRD, Montpellier

Les données : expérience, observation et traitement  
Francis Laloë

Les enjeux de l'information dans le domaine des pêches  
Pierre Chavance

Contrôle de qualité des données. Application à un observatoire socio-économique spatialisé  
Michel Passouant

Recherche d'informations dans un réseau de sources de données scientifiques hétérogènes et autonomes  
Éric Simon

La manipulation de pétaoctets de données en physique des hautes énergies  
Joseph Le Foll

*Coordinateur*, François Le Verge, Ifremer, Brest

Le contrôle qualité dans les centres de données  
François Le Verge, Alain Laponche

Les aspects techniques de la pérennité des données scientifiques

Claude Huc, Danièle Boucon

Video and graphic broadcasting information system for research vessels

*Présentation de l'application SDIV (Système de diffusion d'information et de vidéo) du navire océanographique Thalassa*

Fabrice Lecornu, Armel Rué, Didier Lavoine

Utilisation des techniques avancées : base de données relationnelles, catalogues en ligne www, logiciels expert de contrôle qualité pour l'archivage, la gestion et la diffusion des données océanographiques

Catherine Maillard

Numérisation, transmission, acquisition et traitement de données géophysiques au département Analyse, Surveillance, Environnement du CEA

Pascal Dallot

### 2<sup>e</sup> partie : gestion et valorisation des données

*Coordinateur* : Jean-Michel Kornprobst

Diffusion des données géographiques : valorisation et aspects juridiques

Pierre Peltre

Le partage et la diffusion des données et résultats scientifiques

Dominique Vuillaume



Les données scientifiques : de l'inconduite scientifique à la démarche qualité

Françoise Souyri



Bases de données pour les géosciences : un effort de connaissance et de prospective

Philippe Waldteufel



**Conclusion des débats et synthèse**

Marcel Raffy



La gestion informatique des chroniques en hydrologie

Michel Lang



Gestion et valorisation de données sur l'environnement global, avec l'exemple de Médias-France

Michel Hoepffner, Éliane Cubero-Castan, J.-L. Boichard



### **3<sup>e</sup> partie : aspects juridiques et stratégiques**

*Coordinateur* : Patrick Séchet, IRD, Paris

Les chercheurs peuvent-ils continuer à ignorer le droit ?

Patrick Séchet



Aspects juridiques de la diffusion des données scientifiques

Sébastien Lafargue



Diffusion des données de l'INPI

Bernard Marx



La CNIL et les fichiers de recherche médicale :  
Les nouvelles procédures de formalités dans le secteur  
de la recherche médicale

Jeanne Bossi



## Adresse des auteurs

Jean-Luc **Boichard**, informaticien, Météo-France/Médias, BP 2102, 18, avenue E. Belin, 31401 Toulouse cedex 4.

Jeanne **Bossi**, secteur santé, CNIL, 21, rue St-Guillaume, 75007 Paris.  
e-mail : jbossi@cnil.fr

Danièle **Boucon**, ingénieur CNES, 18, av. Edouard Belin, 34401 Toulouse cedex 4.

Eliane **Cubero-Castan**, informaticienne, Médias-France, BP 2102, 18, avenue Edouard Belin, 31401 Toulouse cedex 4.

Pierre **Chavance** IRD, BP 1386, Dakar, Sénégal.  
e-mail : Pierre.Chavance@ird.sn

Pascal **Dallot**, assistant informatique, CEA/DAM, Analyse, surveillance, environnement, B.P. 12, 91680 Bruyères-le-Châtel.  
e-mail : dallot@dase.bruyeres.cea.fr

Michel **Hoepffner**, hydrologue, IRD-Médias, BP 2102, 18, av. E. Belin, 31401 Toulouse cedex 4.  
e-mail : Michel.Hoepffner@medias.cnes.fr

Claude **Huc**, ingénieur, département Valorisation et gestion des données spatiales, CNES, 18, av. Edouard Belin, 31401 Toulouse cedex 4.  
e-mail : claude.huc@cnes.fr

Jean-Michel **Kornprobst**, professeur université de Nantes, vice-Président de la CS7, ISOMer, Laboratoire de chimie marine, BP 92208, 2, rue de la Houssinière, 44322 Nantes celex 3.  
e-mail : jean-michel.kornprobst@wanadoo.fr

Sébastien **Lafargue**, juriste, Ifremer, Technopolis 40, 155, rue J.J. Rousseau, 92138 Issy-les-Moulineaux.  
e-mail : Sebastien.lafargue@ifremer.fr

Francis **Laloë**, IRD, Halieutique et Écosystèmes Aquatiques, BP 5045, 34032 Montpellier cedex 1.  
e-mail : laloe@mpl.ird.fr

Michel **Lang**, hydrologue, Cemagref, Division hydraulique, 3 bis, quai Chauveau, CP 220, 69009 Lyon cedex.  
e-mail : michel.lang@cemagref.fr

Alain **Laponche**, ingénieur Sismar, Ifremer, centre de Brest, BP 70, 29280 Plouzané.

Didier **Lavoine**, ingénieur réseau, 2 bis, rue R. Le Ricollais, 44000 Nantes.

Fabrice **Lecornu**, ingénieur informaticien, Ifremer, centre de Brest, BP 70, 29280 Plouzané.  
e-mail : Fabrice.Lecornu@ifremer.fr

Joseph **Le Foll**, informaticien, CEA/DSM/DAPNIA, CE Saclay, 91191 Gif-sur-Yvette cedex.  
e-mail : lefoll@hep.saclay.cea.fr

François **Le Verge**, chef du service de la documentation, Ifremer, centre de Brest, BP 70, 29280 Plouzané.  
e-mail : fleverge@ifremer.fr

Catherine **Maillard**, ingénieur de recherche Ifremer, centre de Brest, BP 70, 29280 Plouzané.  
e-mail : Catherine.Maillard@ifremer.fr

Bernard **Marx**, INPI, service DDI, 26 bis, rue de Saint-Pétersbourg, 75008 Paris.

Jean-Pierre **Muller**, pédologue, directeur général de l'IRD, 209-213, rue La Fayette 75480 Paris cedex 10.

Michel **Passouant**, statisticien Cirad, Campus International de Baillarguet, Bât. F, 34398 Montpellier cedex 4.  
e-mail : michel.passouant@cirad.fr

Pierre **Peltre**, géographe, IRD, 32, avenue Henri-Varagnat, 93143 Bondy cedex.  
e-mail : peltre@clarke.bondy.ird.fr

Marcel **Raffy**, professeur, université de Strasbourg, président de la CS7, ULP-CNRS, Parc d'innovation, 5, bd S. Brandt, 67400 Illkirch-Graffenstaden.

Armel **Rué**, ingénieur réseau, Ifremer, centre de Brest, BP 70, 29280 Plouzané.

Patrick **Séchet**, informaticien, IRD, 209-213, rue La Fayette, 75480 Paris cedex 10.  
e-mail : sechet@paris.ird.fr

Éric **Simon**, directeur de recherche en informatique, Inria, BP 105, 78153 Le Chesnay.  
e-mail : eric.simon@inria.fr

Françoise **Souyri**, directeur de recherche, MENRT-CSDR, 5, rue Descartes, Paris cedex 05.  
e-mail : francoise.souyri@dr.education.gouv.fr

Dominique **Vuillaume**, économiste de la santé, Service du partenariat pour le Développement, Inserm, 101, rue de Tolbiac 75654 Paris cedex 13.  
e-mail : vuillaume@tolbiac.inserm.fr

Philippe **Waldteufel**, climatologue, CNRS-IPSL, 10-12, avenue de l'Europe, 78140 Vélizy.  
e-mail : Philippe.Waldteufel@ipsl.uvsq.fr