



Open Access to Biodiversity Scientific Data: A Comparative Study

Melanie Dulong de Rosnay, Andrés Guadamuz

► To cite this version:

Melanie Dulong de Rosnay, Andrés Guadamuz. Open Access to Biodiversity Scientific Data: A Comparative Study. 17th International Consortium on Applied Bioeconomy Research ICABR Conference on Innovation and the Policy for the Bioeconomy, Jun 2013, Ravello, Italy. pp.20. hal-00833480

HAL Id: hal-00833480

<https://hal.science/hal-00833480>

Submitted on 12 Jun 2013

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.

OPEN ACCESS TO BIODIVERSITY SCIENTIFIC DATA: A COMPARATIVE STUDY

MÉLANIE DULONG DE ROSNAY AND ANDRÉS GUADAMUZ

National Centre for Scientific Research (CNRS) Institute for Communication Sciences, France;

University of Sussex, School of Law, UK

melanie.ddd@gmail.com

anduril13@gmail.com



Draft paper prepared for presentation at the 17th ICABR Conference

“INNOVATION AND POLICY FOR THE BIOECONOMY”

Ravello (Italy): June 18 - 21, 2013

This work is licensed under a [Creative Commons Attribution 3.0 Unported License](http://creativecommons.org/licenses/by/3.0/) available at <http://creativecommons.org/licenses/by/3.0/>



ABSTRACT

The field of biodiversity research gathers and crosses collections of environmental data, scientific data, geospatial data and cultural data. All of the available datasets have created a global information pool which can be used to develop further research and downstream bioeconomy innovations. Intellectual Property governance rules vary between proprietary assets and research commons.

The largest datasets are collected, curated, mined, and shared at the national and international level by natural history museums and research institutions which are members of the Global Biodiversity Information Facility (GBIF), an international organization that aims to make available all known information on biodiversity, provide digitization, facilitate interoperability and retrieval of this data.

This paper intends to explore the various legal regimes that apply to biodiversity data. Depending on the data format and the contribution structure, they can be considered as works, public sector information, data, databases, metadata, all submitted to different national and international legal regimes. Besides, contributors apply more or less restrictive terms of use between all rights reserved and open access options.

In order to study the impact of governance choices and practices, the authors have conducted interviews with key stakeholders in several GBIF partner institutions in Europe and Latin America in order to better understand the policy challenges faced by biodiversity researchers who intend to share their information, such as biopiracy, commercial exploitation and scientific attribution for data reuse.

One section of the paper will be dedicated to the study of the legal environment, while the other will focus on contractual practices.

1. INTRODUCTION

This paper provides an overview of the legal and policy conditions governing the contribution and the reuse of biodiversity data within scientific databases. Providers and appropriators of such resources are mostly researchers, natural history museums and botanists. The GBIF, Global Biodiversity Information Facility, is one of the projects ensuring interoperable data can be used by other researchers than those who curated them.

Biodiversity data is at the same time public environmental information, scientific data, taxonomy, metadata, geographic data and cultural heritage works at the crossing between different legislations, natural and knowledge commons.

The study provides an overview of applicable laws, terms of contributions and terms of use. This rely on desk and literature research including applicable law and on interviews of database producers in Europe and in Costa Rica. The objective is to identify the various roles (contributor, editor, publisher, reuser) and the possible reasons to restrict the access to the data. Justifications can be the integrity and the quality of the data and its descriptive metadata, source or attribution of the researcher or the team, the risk of biopiracy related to geolocalization, the role of participative science, and the commercialization of the data output and the risk of international exploitation.

A first observation reveals the lack of harmonization between licensing conditions, hindering interoperability and ease of use for researchers. It will be proposed to minimize contractual uncertainties and develop harmonized contractual clauses or best practises guidelines to facilitate the upstream contribution to the online collections and the downstream reuse and collaborative editing towards a more sustainable information commons.

2. THE CONTRASTING LEGAL ENVIRONMENT

2.1 THE CASE FOR OPEN ACCESS TO SCIENTIFIC RESEARCH

The term “open access” is now widely used to describe publishing methods in peer-reviewed journals through subscription-free models. Open Access was initially developed for publications and journals, but the concept is being extended to data within the Open Data and Open Science movements. This section presents both initiatives and their origins in declarations by scientists and in governmental bodies.

The accepted definition comes from the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities, which states:

“We define open access as a comprehensive source of human knowledge and cultural heritage that has been approved by the scientific community. [...] Open access contributions include original scientific research results, raw data and metadata, source materials, digital representations of pictorial and graphical materials and scholarly multimedia material.”

In similar fashion, the Budapest Open Access Initiative (BOAI) defines open access in light of peer-reviewed and scholarly publications. All of these definitions tend to be very narrow, and geared towards the academic journal market. Philosopher and open access advocate Peter Suber proposes a more open definition, which states that (Suber, 2004):

“Open access” (OA) is free online access. OA literature is not only free of charge to everyone with an internet connection, but free of most copyright and licensing restrictions. OA literature is barrier-free literature produced by removing the price barriers and permission barriers that block access and limit usage of most conventionally published literature, whether in print or online.

This definition tends to be more in line with the sharing ethos that gives birth to open source software (Moody, 2002), and that is the intellectual and ideological parent of open access.

Open access has been gaining momentum around the world. Just recently, the White House released a memorandum in which it directs Federal agencies that fund research to take the adequate steps to make the results available to the public; this includes results published in peer-reviewed scholarly publications that are based on research that directly arises from

Federal funds (White House, 2013).

Similarly, there is growing understanding from governments and funding bodies that the fruits of public research should be shared with the public free of charge. The European Commission recently published a Communication (COM(2012) 401) that states that “information already paid for by the public purse should not be paid for again each time it is accessed or used, and that it should benefit European companies and citizens to the full.” This information should be online free of charge.

Furthermore, some countries have gone further and promised some far-reaching open access policies. In the UK, the Report of the Working Group on Expanding Access to Published Research Findings (Finch Report) has been tasked with advising the government on its policies with regards to scientific research. It has produced detailed policies on the opening up of scientific data that has been obtained with public funds. As a result of the Finch Report, the UK government has made a commitment to implement the recommendations from the Report, and has guaranteed that all future research funded by public money will be available without restrictions anywhere in the world.

The EU Commission Recommendation of 17 July 2012 on access to and preservation of scientific information targets scientific publications and research data that receive public funds, and asks for clear open access policy no later than 6 months after publication, or 12 months for the social sciences.

The Open Access movement is extending to scientific data and databases, with changing social norms in scientific communities, journals publishing agreements and funders or institutions such as the OECD encouraging the deposit of the underlying dataset at the time of the publication of the article. Open scientific data was born among physicists, with a first institutional proposition in 1995 to release Geophysical and Environmental Data as free and unrestricted. The EU has been producing recommendations towards open data for science, without producing any binding requirement mandating the deposit of scientific data, even for

projects funded by the EC research programmes.

Still, the ongoing Horizon 2020 process on open access does not seem to be able to produce a clear open access mandate neither for publications nor for underlying datasets beyond projects funded by EU research programmes. Similarly, the proposal to amend the PSI Directive neither does not include data from research and science, nor data produced with public funding. Here again, all efforts have been left in the hand of individual parties. This situation is leading to scientific data not being accessible not only to the public, but also to their own producers, the scientists and their institutions, the universities and the libraries.

However, despite the great push towards open access for scientific research all across the board, there are two main challenges that are unique to the area of biodiversity that make opening data more difficult than in other disciplines. Firstly, there might be worries about the potential misuse of biodiversity data, particularly by the appropriation of traditional resources by commercial interests, also known as biopiracy. Secondly, the legal protection of biodiversity can be problematic, as we are dealing with various regimes, sometimes contradictory to one another.

2.2 THE CONFLICT WITH BIOPIRACY

While open access is generally seen as a positive, in the areas of biotechnology in general, and biodiversity specifically, there could be some concerns about the manner in which information is made available to be shared to a wider audience, particularly if the sharing is not accompanied by some sort of legal protection.

For example, during the chase to sequence the human genome, the Human Genome Sequencing Consortium placed large amounts of data in the public domain; the idea being that sharing this information would lead to a faster decoding process as more data online would translate into more researchers benefiting from the released information. However, former members of the Consortium tried to take advantage of the released data and set up a competing entity that

would attempt to commercialise the results (Guadamuz, 2006). While the commercial efforts eventually failed, the fears of misuse might continue.

Advances in biotechnology in the last decades have greatly increased the economic value of biodiversity. It is a fact that warmer zones in the world contain greater range of biodiversity, and most of the areas with the greatest biodiversity can be found in the developing world. The cataloguing of such biodiversity has begun in recent years, prompting the discovery of new resources, that is, plants and animals that may hold immense medicinal powers. Coupled with this, many indigenous people have been using some of these traditional medicines for generations, and among these, there may be medicines that could be used all around the world. The search of these potentially profitable plants or animals has prompted several pharmaceutical companies to spend relatively small amounts of money in looking for possible sources of new medicines in the biodiverse rich areas located in developing countries. This has led to the possibility that resources from developing countries could make their way to developed nations, where they would be converted into medicines, patented and then shipped back and sold at high prices in the same countries where it was originated. It can also create problems for countries wanting to export medicines or plant varieties derived from traditional knowledge and based on plants that can be found in those countries, as they could find that these innovations are already protected in developed countries. Some commentators have named this phenomenon "biopiracy" (Shiva, 1997).

The term "biopiracy" is a vague, political term used to describe actions by corporations from developed countries which take unfair advantage of genetic resources or traditional knowledge of developing countries. These actions may either be illegal, or legal (but considered morally wrong or unfair). Cases have been referred to as "biopiracy" in which a patent was wrongfully granted for inventions that were not inventive or novel. This form of "biopiracy" is not a problem of the patent system in itself, but rather of the examination procedure and the accessibility of traditional knowledge to examiners.

The awareness of biopiracy has increased in recent years, particularly the bio-prospecting of plants and remedies in China and India. There are numerous examples of biopiracy in India, but perhaps one of the most indicative cases is that of the plant *phyllanthus niruri*, used in India for many years for the treatment of jaundice and liver deficiencies. In 1985, the Fox Chase Cancer Center in the United States applied for a patent for the extract of this plant as a cure for hepatitis B, which was granted in US patent 4673575. Shiva (Shiva, 2001) argues that this patent is the exact same use that has been applied in traditional Ayurveda medicine, which would seemingly invalidate the patent because it lacks the novelty requirement. In 1988, the same centre applied for another patent for the use of the same plant as an antiretroviral medicine, which they also obtained.

Another well-publicised case is that of a patent awarded for basmati rice in the United States. Basmati rice is an increasingly popular variety of rice that has been cultivated and used in India and Pakistan for generations. In 1997, a Texan company called RiceTec Inc. was issued with patent 5663484. The Indian government spearheaded a campaign to have the patent cancelled. Eventually, most of the claims from this patent were either dropped by RiceTec or cancelled by the USPTO for the existence of prior-art in the application. The patent remains in a very limited form that protects only certain specific basmati varieties bred by the company.

While the above examples are mostly about the misappropriation of traditional knowledge through patents, there are other examples of biopiracy where it is alleged that there is commercialisation of biodiversity. The most famous case is that of Costa Rica. In 1991, the Costa Rican Instituto Nacional de Biodiversidad (INBio), a semi-private, non-profit, scientific organization, signed with the multinational pharmaceutical company Merck and agreement by which INBio would supply Merck with samples among the plants, insects and microorganisms collected from Costa Rica's protected forests in exchange for equipment

and some funds (Zebich-Knos, 1997). While this may seem like an excellent example of benefit sharing, some have criticised it precisely because it commercialises biodiversity, and Costa Rica sold it cheaply (Gulati, 2001).

However, it must be pointed out that some consider that there is no such thing as biopiracy (Chen, 2006). Many of the examples cited as evidence have not been commercialized, or the patents have been repealed. Similarly, those who criticise the concept point that there is little evidence of the widespread plunder of biological resources from developing countries. Nevertheless, the debate continues, and the prospective of the existence of biopiracy and bioprospecting might place a dampener on the advance of open access. Why? Because if data from indigenous knowledge and developing countries is made available under open access conditions, pharmaceutical companies will be able to exploit them without compensation for the local communities.

2.3 THE LEGAL PROTECTION OF BIODIVERSITY DATA

According to the article 2 of the Convention on Biodiversity (CBD), biodiversity is understood as the variability among living organisms from all sources, including terrestrial, marine and other aquatic ecosystems and the ecological complexities of which they are part; this includes diversity within species, between species and of ecosystems.

The scientific study of such diverse materials makes for an incredibly complex scenario of legal regimes involved. This is because of the nature of scientific research itself, but also has to do with the type of research that is conducted. In general, biodiversity investigation involves a combination of methods that can prove tricky for the law. We have collection of samples by researchers (and sometimes by members of the public), then the storage and cataloguing of samples, taking of pictures, *in situ* notes, creation of information databases, bioprospecting,¹ publication of results, display of data through websites, archiving in institutional repositories,

¹ Which can be defined as the collecting samples of biodiversity (and sometimes also related indigenous knowledge) in order to develop commercial products, eg. drugs and crops.

patent applications, etc. The above does not fall easily into existing categories of intellectual property protection and may trigger several (e.g. copyright, database, patents, trademark, plant varieties, trade secrets, environmental and geographic data, public sector information), and in some instances, it might not even be protected.

For example, the notes and pictures taken during a collection of samples would be protected by copyright law as literary and artistic works respectively, while any samples or specimens themselves would not be protected. Observation results and comments made on specimen reproductions of a plant or an animal are indispensable to the understanding of the primary data that is of little value without metadata. However, copyright law could be providing a stronger protection to the photography than to the underlying data to the extent data and taxonomy are not considered as creative work. This discrepancy reveals an epistemological gap between copyright law and scientific effort conceptions of a creative or original effort, the threshold of protection.

There is also a difference whether data is considered as cultural, public, geographic or environmental data. Biodiversity data can be considered as a cultural data to the extent databases can be produced and held by a museum of natural history and therefore outside of the scope of the 2001 Directive on reuse for Public Sector Information. It can be considered as public sector information collected by a public body and submitted to an obligation or use for any purpose, including commercial or research under the scope of this Directive. Finally, part of the biodiversity data can fall under the scope of provisions governing geographic information, in the case of geolocation data associated to a specimen and thus governed by the INSPIRE Directive establishing an Infrastructure for Spatial Information in the European Community providing a useful framework for data sharing, in particular giving public authorities “the technical possibility to link their spatial data sets”, but offering the possibility to restrict access for reasons related to Intellectual Property Rights, without mentioning whether it is copyright or patent, or for “the protection of the environment to which such information relates, such as the

location of rare species”, potentially cancelling the effect for biodiversity data. Finally, the Directive 2003/4/EC of the European Parliament and of the Council of 28 January 2003 on public access to environmental information and the Convention of Aarhus guarantee the right of access, also subjected to Intellectual Property Rights and “protection of the environment to which such information relates, such as the location of rare species”. The fact that public authorities provide the right of access does not convey a right of reuse, which may be necessary for data mining or to otherwise process databases electronically.

Moreover, some aspects of biodiversity research might be protected in one jurisdiction, but not in another. For example, data contained in a scientific database can be protected by copyright in some jurisdictions as a literary work. For example, in the UK, Section 3A of the Copyright, Designs and Patents Act 1988 (CDPA), defines a database as a collection of independent works which "are arranged in a systematic or methodical way", and "are individually accessible by electronic or other means". However, the threshold of originality in a database is quite high, just as it happens in the United States. To complicate matters, the European Union has implemented a sui generis right arising from the European Database Directive (96/9/EC). The database right is an exclusive right given to the maker of a database, which is defined as a collection of independent works, data or other materials that are arranged in a systematic or methodical way, and are individually accessible by electronic or other means. It is important to point out that the database right exists regardless of the existence of copyright protection in the database, as the exclusive rights given to the database owner are separate to those arising from copyright.

The following table can help to illustrate the complexity of the legal protection of biodiversity data:

Table 1: Legal protection of biodiversity data:

Copyright	Patents	Database right (EU)	Other/No protection
<ul style="list-style-type: none"> - Notes - Pictures - Metadata - Publications 	<ul style="list-style-type: none"> - Bioprospecting - Processes - Methods - Software 	<ul style="list-style-type: none"> - Data - Databases - Thesaurus - Ontologies 	<ul style="list-style-type: none"> - Brands - Materials - Specimens - Plant varieties

<ul style="list-style-type: none"> - Reports - Data - Software - Compilation 			<ul style="list-style-type: none"> - Genetic banks - Designs - Geolocation information
--	--	--	---

So from the start we see some of the challenges faced by parties interested in opening access to biodiversity. This uncertainty makes it difficult to understand what rights may apply to what parts of databases, jeopardizing interoperability, fusion and machine processing of large amounts of data.

Besides the legal complexity and discrepancy highlighted, there are other conflicting legal regimes. Originally, plant genetic resources were considered as common heritage of mankind, meaning that the resources were freely available and accessible to all without permission. This principle was gradually undermined as patents covering biological material were increasingly granted by developed countries. In reaction, developing nations strived to reinforce the principle of common heritage of mankind in relation to genetic resources, which led to the negotiation of a treaty that demanded national sovereign rights over genetic resources and pushed toward benefit sharing provisions. The result is the 1992 Convention of Biological Diversity. It is the first international treaty to address conservation and sustainable use of all biological diversity worldwide. Its objectives are (cf. Art. 1 CBD):

1. The conservation of biological diversity,
2. The sustainable use of its components,
3. The fair and equitable sharing of the benefits arising out of the utilisation of genetic resources.

The CBD asserts sovereign rights of nations over their national resources (Art. 3, 15 No. 1), notes that access to genetic resources should be subject to prior informed consent (Art. 15 No. 5), and determines that results of R&D and the benefits of commercialisation and utilisation of genetic resources should be shared fair and equitably on mutually agreed terms (Art. 15 No. 7).

As mentioned above, one of the objectives of the CBD is "the fair and equitable sharing of the

benefits arising out of the utilisation of genetic resources". Following this important principle, a Working Group was established in 2004 to negotiate an international regime on access to genetic resources and benefit-sharing with the aim of adopting an instrument to effectively implement the CPD provisions. A protocol was agreed at in Nagoya, Japan, in October 2010. The Nagoya Protocol on Access and Benefit-sharing continues with the line set out in the CBD by stating in Art. 5 that "benefits arising from the utilization of genetic resources as well as subsequent applications and commercialization shall be shared in a fair and equitable way with the Party providing such resources that is the country of origin of such resources or a Party that has acquired the genetic resources in accordance with the Convention."

Besides those international obligations, some countries have passed legislation that makes an obligation to share benefits arising from biodiversity research and bioprospecting. The most cited example is the 1998 Biodiversity Law of Costa Rica (Law 7788), which creates several obligations for those involved in the act of bioprospecting in its territory. Firstly, the law establishes a requirement for anyone engaged in the act of biodiversity research for any purpose to register to a specially created commission, which will then grant permission to engage in bioprospecting. Secondly, the law creates several useful figures that take local communities into consideration, such as prior informed consent, technology transfer and the equitable distribution of benefits. The law recognises the intellectual property where applicable, but established licenses for communities, as well as creating a sui generis "community right". In short, the Law 7788 is perhaps one of the most comprehensive pieces of legislation that not only protect biodiversity, but that seems to bridge the dichotomy between the CBD and intellectual property protection.

But how well does it work in practice?

3. DATABASES TERMS OF USE

After the identification of applicable law, the article will analyse the contractual provisions which may apply to databases containing biodiversity scientific data. Indeed, in addition to the framework of the law, databases producers have the possibility to apply contracts, otherwise named licences or terms of uses to the websites hosting the data they make available to the public. Various options exist, from the most restrictive provisions to the most liberal terms of use based on the principles of Open Access. One of the key clauses of those contracts is the one regulating attribution of the author, which is the legal counterpart for the scientific norm of citation.

3.1 CONTRACTUAL RESTRICTIONS TO ACCESS AND REUSE

Copyright and other related rights allow producers of data to exclude users from accessing or reusing them. While some decide to not fully apply their rights, or partially waive some to embrace Open Access as it will be described in the next section, it is a widespread practice to apply a copyright by default, or to express confusing terms of use. For instance, the biodiversity database of the Botanical Garden of the City of Geneva (<http://www.ville-ge.ch/cjb/bd.php#catherb>) reserves rights for education and research, de facto excluding companies working independently or which can be partners of research projects with public institutions. Its terms of use are contradictory, on the one hand they mention that data can be freely reused, but on the other hand that a request must be addressed for reuse (http://www.ville-ge.ch/musinfo/bd/cjb/chg/condition_utilisation.php?lang=fr).

Another common usage is to apply directly restrictive terms of use and exclude commercial use or any usage other than private consultation, such as the New Mexico Biodiversity Collections Consortium (NMBCC) Restrictions on Use (<http://nmbiodiversity.org/fineprint.php>) or French Plantnet (<http://ee.plantnet-project.org/mentions>).

Knowledge commons framework has been applied to biodiversity data (Louafi & al, 2012;) prior

research established that contractual restrictions is damaging for science and impeding data processing. Guidelines to share scientific data have been developed by the Open Knowledge Foundation (Open Knowledge Definition), Science Commons (Science Commons, 2011) and GBIF, the Global Biodiversity Information Facility (Moritz & al, 2011). All agree to avoid contractual and technical restrictions to download and reuse the data, and put an emphasis on reusability. Several licensing framework and tools are available to reach that objective.

3.2 OPEN ACCESS LICENCES FOR SCIENTIFIC DATA

Some institutions choose to freely disseminate their scientific data and allow the broadest range of access and reuse. They have several choices of standard licences, the Creative Commons licencing set being the most common choice.

Databases producers can also write their own terms of use and declare they do not assert Intellectual Property Rights such as GBIF (<http://data.gbif.org/tutorial/datasharingagreement>). GBIF is embracing a wide Open Access policy and is trying to convince the members to follow this example. Searchable information which can be accessed to and reused among disciplines is required to ease the work of scientists and conservation efforts, which should not be delayed by “old-fashioned practices of handling scientific data” (Fonseca, Benson, 2003).

Only full Open Access terms as close as possible to the public domain will allow unrestricted data mining. Otherwise, partial Open Access and the use of different licences will raise legal interoperability issues (Dulong de Rosnay, 2010). Some contributors consider the absence of legal requirement to attribute as an impediment. This concern led to the withdrawal of a few institutions members of the project “OpenUp !” Opening up the Natural History Heritage for Europeana after the change of Europeana foundation to public domain conditions for the contributors to metadata (Europeana data exchange agreement, <http://pro.europeana.eu/web/guest/data-exchange-agreement>). Others consider that the scientific norm is sufficient to regulate proper attribution and that a contractual requirement

could lead to licensing stacks in the case of aggregation of large numbers of databases (Science Commons, 2011; Hagedorn & al, 2011).

3.3 PARTICIPATIVE PLATFORMS FOR DATA PUBLISHING

The main challenge to Open Data is the attribution of the contributor which gives an incentive to participate for the researcher, but also credibility to the data produced. The validation by experts of a classification and the curation of taxonomy issues also provide a quality-control mechanism. Such a peer-review mechanism has been developed by WoRMS, the World Register of Marine Species (Costello & al, 2013). GBIF developed a Data Publishing Framework with a data citation mechanism, making it easy to attribute correctly the researcher who created or added value to the data (Chavan, V. S., & Ingwersen, P., 2009) and recommends to establish a “data citation service” facilitating deep-data citation” (Moritz & al, 2011).

The connection between the publication of an article and the provision of the underlying data is another path that will provide an incentive for authors to distribute their data under Open access conditions. The journal in plant systematics PhytoKeys provides methods for data publication, linking data to metadata and indexation to facilitate data mining (Penev & al, 2010). Next to these efforts, the following step would be for universities and research institutions to integrate the contribution to a database and its citations in the evaluation of the careers.

4. CONCLUSIONS AND NEXT STEPS

Biodiversity data can be subjected to different conflicting legal provisions and a clarification of their legal status is required. Several institutions are pioneering work related to Open Access to data, including automated mechanisms for citation. These efforts should be developed and adopted across databases in order to replace terms of use that makes it impossible to perform data mining and reuse the data. Future work requires a more systematic analysis of terms of use and interviews of stakeholders and database producers in order to develop recommendations and contractual mechanisms.

5. ACKNOWLEDGEMENTS

This research has been funded in 2012 by the Interdisciplinary Programme on Communication Sciences of the French National Centre for Scientific Research. Clément Guillou of Ecole Centrale d'Electronique in Paris was helpful in retrieving some databases and their terms of use. The authors would also like to thank researchers who were interviewed, all mistakes remaining ours and not implying endorsement: Anne-Sophie Archambeau, Donald Hobern, María Mora and Vishwas Chavan at GBIF France, GBIF Secretariat and Inbio, GBIF Costa Rica; Nicolas Boeglin at University of Costa Rica, Silvia Rodriguez Fernandez from the Universidad Nacional in Costa Rica.

6. REFERENCES

6.1 LEGAL TEXTS AND REPORTS

Accessibility, sustainability, excellence: how to expand access to research publications. Report of the Working Group on Expanding Access to Published Research Findings. United Kingdom <http://www.researchinfonet.org/publish/finch/>

Agreement on Trade-Related Aspects of Intellectual Property (TRIPS): only focus on the aspects related to plant genetic resources

Berlin Declaration on Open Access.

Convention on Biological Diversity (CBD)

Costa Rica, *Ley de Biodiversidad* N° 7788, 1998.

Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases.

European Commission. (2012) *Towards better access to scientific information: Boosting the benefits of public investments in research*, COM(2012) 401.

FAO Global Plan of Action (GPA) for the Conservation and Sustainable Utilization of Plant Genetic Resources for Food and Agriculture;

Final Report of the High-Level Group on Scientific Data to the European Commission, *Riding the Wave: How Europe can gain from the rising tide of scientific data*, 2010.

International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA, also known as the International Sees Treaty)

Macdonald, A et al. European Study Towards a European eInfrastructure for eScience digital repositories, study for the European Commission, DG Information Society and Media Unit F – GÉANT and e-Infrastructure, e-SciDR, Project reference no: 2006 S88-092641, 2008.

OECD Principles and Guidelines for Access to Research Data from Public Funding, 2004, Committee on Geophysical and Environmental Data - National Research Council, *On the Full and Open Exchange of Scientific Data*, 1995.

Open Knowledge Definition. <http://opendefinition.org/okd/>

Proposal for a Regulation of the European Parliament and Council establishing Horizon 2020 - the Framework Programme for Research and Innovation (2014-20), 30 November 2011.

Science Commons. (2011). Science Commons Protocol for implementing open access data. <http://sciencecommons.org/projects/publishing/open-access-data-protocol/>.

The Joint Data Archiving Policy requires as a condition to be published in several evolution journals to deposit underlying data in a repository, <http://datadryad.org/jdap>

The Wellcome Trust, the Canadian Institutes of Health Research (CIHR) policy, <http://www.cihr-irsc.gc.ca/e/32326.html>.

UPOV Convention

White House. (2013). *Increasing Access to the Results of Federally Funded Scientific Research*,

http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf.

6.2 ARTICLES AND BOOKS

Burk, D. (2002). Open Source Genomics. *Boston University Journal of Science and Technology Law*, 8, 254.

Boettiger, S., & Burk, D. L. (2004). Open Source Patenting. *Journal of International Biotechnology Law (JIBL)*, 1(6), 11.

Chen, J. M. (2006). "There's No Such Thing as Biopiracy ... And It's a Good Thing Too", *McGeorge Law Review* 37, 16.

Chavan, V. S., & Ingwersen, P. (2009). "Towards a data publishing framework for primary biodiversity data: challenges and potentials for the biodiversity informatics community". *BMC Bioinformatics*, 10 Suppl 14, S2.

Chavan, V., & Penev, L. (2011). "The data paper: a mechanism to incentivize data publishing in biodiversity science". *BMC Bioinformatics*, 12 Suppl 15, S2.

Costello, M. J., et al. (2013). "Global Coordination and Standardisation in Marine Biodiversity through the World Register of Marine Species (WoRMS) and Related Databases". [doi:10.1371/journal.pone.0051629]. *PLoS ONE*, 8(1), e51629.

Cukier, K. (2003). "Open Source Biotech: Can a non-proprietary approach to intellectual property work in the life sciences?". *The Acumen Journal of Life Sciences*, 1(3).

Dedeurwaerdere T., Stromberg, P. M. and Pascual, U. (2012). Social Motivations and Incentives in Ex Situ Conservation of Microbial Genetic Resources, in Dulong de Rosnay M. and De Martin JC. (eds), *The Digital Public Domain: Foundations for an Open Culture*, OpenBook Publishers, Cambridge, UK.

Drahos, P. (2000). Indigenous Knowledge, Intellectual Property and Biopiracy: Is a Global Bio-Collecting Society the Answer? *European Intellectual Property Review*, 22(Part 6), 245.

Dulong de Rosnay M. (2010). Creative Commons Licenses Legal Pitfalls: Incompatibilities and Solutions, *study of the Institute for Information Law of the University of Amsterdam*.

Dutfield, G. (2003). *Intellectual Property Rights And The Life Science Industries: A 20th Century History*. Aldershot: Ashgate.

Fonseca, G., & Benson, P. J. (2003). "Biodiversity Conservation Demands Open Access". *PLoS Biol*, 1(2), E46.

Gibson, J. (2008). *Patenting Lives: Life Patents, Culture and Development*. Aldershot: Ashgate.

Guadamuz Gonzalez, A. (2006). Open Science: Open Source Licences for Scientific Research. *North Carolina Journal of Law and Technology*, 7(2), 321.

Goel, R. (2008). "Protection and Conservation - TRIPs and CBD: a Way Forward", *Journal of Intellectual Property Law & Practice* 3, 334.

Gulati, C. (2001). "The "Tragedy of the Commons" in Plant Genetic Resources: The Need for a New International Regime Centered Around an International Biotechnology Patent Office", *Yale Human Rights & Development Law Journal* 4, 63.

Hagedorn, G., Mietchen, D., Morris, R. A., Agosti, D., Penev, L., Berendsohn, W. G., et al. (2011). Creative Commons licenses and the non-commercial condition: Implications for the re-use of biodiversity information. *Zookeys*(150), 127-149.

- Herr, R. E. (2008). *Is The Sui Generis Right a Failed Experiment: A Legal and Theoretical Exploration of How to Regulate Unoriginal Database Contents and Possible Suggestions for Reform*. Copenhagen: DJØF Pub.
- Hope, J. (2007). *Biobazaar: The Open Source Revolution And Biotechnology*. Cambridge, Mass.: Harvard University Press.
- Karjala, D. S. (2006). "Biotech Patents and Indigenous Peoples", *Minnesota Journal of Law, Science & Technology* 7, 483-527.
- Louafi S, Arnaud E, Barthelemy D, Noyer J-L, Pham J-L. (2012). "Value, Norms And Practices In Plant Biodiversity-Based Research And Innovation Commons". *First thematic conference on the knowledge commons. Governing pooled knowledge resources: Building institutions for sustainable scientific, cultural and genetic resource commons*. 12-14th September 2012, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.
- Moody, G. (2002). *Rebel code: Linux and the open source revolution*. London: Penguin.
- Penev, L., Kress, W. J., Knapp, S., Li, D. Z., & Renner, S. (2010). Fast, linked, and open - the future of taxonomic publishing for plants: launching the journal PhytoKeys. *PhytoKeys*(1), 1-14.
- Shiva, V. (2001). *Protect Or Plunder? Understanding Intellectual Property Rights*. London: Zed Books.
- Shiva, V. (1997). *Biopiracy: The Plunder Of Nature And Knowledge*. Boston, MA: South End Press.
- Silvertown, J. (2010). Taxonomy: include social networking. *Nature*, 467, 788.
- Spinney, L. (1998). *Biotechnology in Crops: Issues for the developing world*: Oxfam GB.
- Stenton, G. (2004). Biopiracy within the Pharmaceutical Industry: A Stark Illustration of how Abusive, Manipulative and Perverse the Patenting Prices can be towards Countries of the South. *European Intellectual Property Review*, 26(Part 1), 17.
- Suber, P. (2004). *Creating an Intellectual Commons through Open Access*, Report: http://dlc.dlib.indiana.edu/dlc/bitstream/handle/10535/4445/Suber_Creating_041004.pdf.
- Westerlund, L. (2002). *Biotech Patents: Equivalence and Exclusions Under European and U.S. Patent Law*. The Hague; New York: Kluwer Law International.
- Zebich-Knos, M. (1997). "Preserving Biodiversity in Costa Rica: The Case of the Merck-Inbio Agreement". *Journal of Environment and Development*, 6(2), 180.