

Micro B3: Marine Microbial Biodiversity, Bioinformatics and Biotechnology

Call: Joint Call OCEAN.2011-2: Marine microbial diversity – new insights into marine ecosystems functioning and its biotechnological potential

Grant agreement no.: 287589

Deliverable 8.2: IP Model Agreements for Pre-Competitive Access to Large-Scale Microbial Genomic Research Databases

Lead participant: UCL

Partners contributing: UniHB, CIESM, IUCN

[Please see the document control sheet]

Due date of deliverable: 31 December 2012

Actual submission date: 23 January 2013

Nature of deliverable: Report

Status (draft/final): Final

Project funded by the European Commission within the Seventh Framework Programme	
Dissemination Level	
PU Public	X
PP Restricted to other programme participants	
RE Restricted to a group specified by the consortium (including the Commission Services)	
CO Confidential, only for members of the consortium (including the Commission Services)	

Document control sheet

Version	Author	Date	Status	Reviewer
V.1	Arul George Scaria, Arianna Broggiato and Tom Dedeurwaerdere	17/12/2012	Draft	Tom Dedeurwaerdere [Contributions from WP8 members on these issues through (1) kick off meeting in Bremen, (2) WP8 Skype meetings and (3) a workshop in Bonn in November 2012] and through circulation and discussion of draft documents for the model Agreement on Access to Marine Microorganisms and Benefit Sharing by email.
V.2	Arul George Scaria, Arianna Broggiato and Tom Dedeurwaerdere		Final	

Table of Contents

Section 1. Introduction.....	3
Section 2. Relevant legal and policy framework under international and EU law	4
2.1 Initiatives for early release of genomic reference data.....	4
2.1.1 Bermuda Principles of 1996.....	5
2.1.2 Fort Lauderdale Principles of 2003.....	5
2.1.3 Toronto Statement of 2009	7
2.1.4 EC recommendation C/2008/1329.....	8
2.1.5 OECD guidelines for access to research data from public funding	8
2.1.6 Implementation of the data policies and some of the important challenges	9
2.2 Marine data and UNCLOS	11
2.3 Copyright and database rights	13
2.4 Data License Agreements	16
2.5 Practical implications	17
Section 3. Two case studies	17
3.1 Case study 1 – Database of British Oceanographic Data Center	17
3.2 Case study 2: BONUS EEIG Data Policy	18
3.3 Practical Implications	20
Section 4. Model Agreement and incorporation of the core principles in the ABS Agreement	20
Section 5. Conclusion and Future Directions	26
Annex 1	28

Section 1. Introduction

The primary objective of deliverable 8.2 is to develop sector specific model agreements for pre-competitive access to microbial research data generated in marine bioprospecting missions. As stated in the project document, the model agreements are developed based on an adaptation of the Fort Lauderdale Agreement for fair use of data of large-scale genome sequencing projects. This is with the objective of enabling the pooling of data from large-scale community resource projects and fostering rapid pre-publication data releases, thereby creating a robust public domain for research and innovation.

The three specific issues to be addressed in implementing these general objectives are ensuring (1) appropriate rules for attribution / citation of data, (2) data interoperability, and (3) automated data integration. The participants under Work Package 8.2 have been discussing this topic in several project related meetings, including the kick off meeting in Bremen, WP8 Skype meetings and a workshop in Bonn in November 2012. The circulation and discussion of the draft documents relating to the Model Agreement on Access to Marine Microorganisms and Benefit Sharing (developed in conjunction with Deliverable 8.1), wherein specific data related provisions are also included, further helped to complete the objectives within the stipulated time frame. This report summarises the work undertaken as part of this Deliverable. This report is structured as follows: Section 2 provides a review of the relevant legal and policy framework, both at the international and European levels. Section 3 provides two interesting case studies that have been analysed by the project team and which illustrate the open access efforts with regard to large scale databases. Section 4 provides one of the model agreements developed under Deliverable 8.2 and also discusses how the key aspects of the model agreement were also included in the Model Agreement on Access to Marine Microorganisms and Benefit Sharing (“herein after referred to as “model ABS Agreement”) developed in conjunction with Deliverable 8.1 of the project dealing with the material transfer issues. Section 5 concludes the report and provides some future directions, in particular for the next steps within the project, which is the development of the more specific MICRO B3 data policy guidelines based on this work.

Section 2. Relevant legal and policy framework under international and EU law

The developments in sequencing technologies have produced unprecedented challenges with regard to management, use and storage of data. The exceptional quantities of data developed through the new technologies have also raised concerns regarding the optimal utilisation of those data for the wider progress of science. These challenges also make it highly important to ensure that any model agreement developed as part of this project deliverable is not only legally valid, but also promote the optimal utilisation of data for fostering the progress of science.

One of the most important steps in the process of developing any such model agreement is analysis of the legal and policy aspects governing the intended area of operation of the model agreement. This section is intended to provide an overview of the legal and policy framework governing pre-competitive access to microbial research data generated in marine bioprospecting missions.

2.1 Initiatives for early release of genomic reference data

While many individual investigators are tempted to keep the data produced by them as secret for a long period of time, the social costs involved in such tendencies are very high. Broader and more rapid access to data plays a major role in the innovation process of any country/region and hence many research funding institutions and leaders from the science communities have attempted to address this challenge through different approaches.¹

One of the most important steps in this regard was the evolution of data policy initiatives on early release and sharing of data from large-scale biological and genomic research projects. Most of these policies mandated the researchers to deposit their data in open public repositories. This was intended to make the cumulative genomic inputs available to the scientific community as a whole. One may also observe that such community evolved data

¹ See, for example, European Commission Press Release (Brussels, 17 July 2012), 'Scientific data: open access to research results will boost Europe's innovation capacity', http://europa.eu/rapid/press-release_IP-12-790_en.htm (accessed 14 December 2012).

sharing policies are today playing a far more important role than legislative efforts in the area. This section highlights some of the most important initiatives in this regard.²

2.1.1 Bermuda Principles of 1996

The first major initiative in the field is commonly referred to as the *Bermuda Principles of 1996*.³ This initiative advocated two core principles and they were - (1) primary genomic sequence data should be in public domain and (2) primary genomic sequence data should be rapidly released.⁴ The *Bermuda Principles of 1996* advocated the release of data within twenty four hours from the time of generation of data, when the sequence assembly exceeds one kilobase (and later to 2 kb) pairs.⁵

It was also agreed that the *Bermuda Principles* applied to “*all human genomic sequence generated by large-scale sequencing centres, funded for the public good...to prevent such centres establishing a privileged position in the exploitation and control of human sequence information*”.⁶ The *Bermuda Principles* were widely followed and they are still considered as the foundation of a successful system for achieving rapid and open data.⁷

2.1.2 Fort Lauderdale Principles of 2003

The next major initiative in this regard came from a meeting at Fort Lauderdale (Florida), in the year 2003, and this initiative is commonly referred to as “*Fort Lauderdale Principles of 2003*”.⁸ The meeting was sponsored by the Wellcome Trust and it reaffirmed the *Bermuda*

² The discussion in this sub-section is a summarised version of the discussion in Jerome H. Reichman, Tom Dedeurwaerdere and Paul F. Uhler, *Global Intellectual Property Strategies for the Microbial Research Commons : Governing Digitally Integrated Genetic Resources, Data and Literature* (Cambridge University Press, forthcoming 2013), 256-322. For detailed discussion, please refer to the original source.

³ For a summary of the principles agreed at the first international strategy meeting on human genome sequencing (commonly referred to as Bermuda Principles of 1996), see http://www.ornl.gov/sci/techresources/Human_Genome/research/bermuda.shtml (accessed 2 December 2012). For a historical analysis of the principles, see Eliot Marshall, ‘Bermuda Rules: Community Spirit, With Teeth’, *Science*, Vol. 291, Issue 5507, 1192.

⁴ Ibid.

⁵ Ibid.

⁶ Ibid.

⁷ The Wellcome Trust, *Sharing Data from Large-Scale Biological Research Projects: A System of Tripartite Responsibility*, 4, <http://www.genome.gov/Pages/Research/WellcomeReport0303.pdf> (accessed 2 December 2012).

⁸ For details of this initiative, Ibid., 2-6.

Principles.⁹ The meeting made major efforts to extend the principle of rapid pre-publication release to “other types of data from other large-scale production centres specifically established as ‘community resource projects’”.¹⁰ The *Fort Lauderdale Principles* also insisted that “[t]here should be no restrictions on the use of the data”.¹¹

But taking into consideration the possibility that the broadened system of pre-publication data release might conflict with the basic scientific incentive of researchers in publishing the first analysis of one's own data, the *Fort Lauderdale Principles* recommended a tripartite sharing of responsibilities among sequence producers, sequence users and the funding agencies.¹² It specifically affirmed that “[t]he contributions and interests of the large-scale data producers should be recognized and respected by the users of the data, and the ability of the production centres to analyse and publish their own data should be supported by their funding agencies”.¹³ They also made specific recommendations for users in this regard, which included duties of citation, acknowledgement and the obligation to “[a]ct responsibly to promote the highest standards of respect for the scientific contributions of others”.¹⁴

The meeting also noted that many other large data sets will be produced in the near future as community resources and the development of effective systems for achieving the objectives of the community resource concept should be an integral component of the planning and development of such new community resources.¹⁵ The meeting also noted that there are many other valuable data sets which are not technically considered as “community resource projects”, but where resource generation is not the primary goal.¹⁶ According to the meeting, contribution of the data to the public domain as a resource is more on a voluntary basis in such cases and “[i]ncentives should be developed by the scientific community to support the voluntary release of such data prior to publication, by appropriately recognizing and protecting the interests of scientists who wish to share such pre-publication data with the community.”¹⁷

⁹ Ibid., 2.

¹⁰ Ibid. The *Fort Lauderdale Principles* defines ‘community resource project’ as a research project specifically devised and implemented to create a set of data, reagents or other material whose primary utility will be as a resource for the broad scientific community.

¹¹ Ibid., 4.

¹² Ibid., 3-4.

¹³ Ibid.

¹⁴ Ibid.

¹⁵ Ibid., 5.

¹⁶ Ibid., 5.

¹⁷ Id., at 5.

2.1.3 Toronto Statement of 2009

The strong wordings in the Fort Lauderdale Principles in support of the interests of data users as compared to that of data producers (at least in the case of “community resource projects”) worried many data producers about effective preservation of their rights and their publication prospects. Many of these fears were softened to a great extent at a subsequent data release workshop in Toronto in May 2009, organised by Genome Canada and other funding agencies.¹⁸

The resulting *Toronto Statement* built upon the *Fort Lauderdale Principles* and further refined them. Most importantly, it endorsed the principle of extending early data release policies beyond genomics and proteomics studies to other data sets including chemical structure, metabolomics and RNA interference data sets.¹⁹ It even suggested extension of the principles to annotated clinical resources in appropriate cases.²⁰ The *Statement* also emphasized the importance of simultaneously releasing metadata such as environmental or experimental conditions and phenotypes for enabling data users to fully exploit the data.²¹

The *Toronto Statement* further clarified the duties of data users to respect the rights and also the publishing expectations of data producers. It expressly obliged the users to respect the embargo periods in favour of data producers for first publication, that would “*ideally expire within one year*”.²² Simultaneously, the data producers were also urged to provide marker papers associated with their database entries.²³ This was primarily intended to enable citations and tracking of the usage of early released data.²⁴

The *Toronto Statement* also expressly distinguished between the large scale reference data sets, which were suitable for the application of early release policies, from the hypothesis driven datasets, where a delay in release was appropriate until the publication of relevant findings.²⁵ The *Toronto Statement* also attempted to address some of the quality issues with early release of datasets, with a caution to editors and reviewers to look for possible sources

¹⁸ For an excellent summary of the workshop outcomes, see ‘Prepublication Data Sharing’, *Nature*, (2009) Vol. 461, Issue 10, 168, <http://www.nature.com/nature/journal/v461/n7261/full/461168a.html> (accessed 3 December 2012).

¹⁹ *Ibid.*, 168.

²⁰ *Ibid.*, 168.

²¹ *Ibid.*, 169.

²² *Ibid.*, 170.

²³ *Ibid.*, 170.

²⁴ *Ibid.*

²⁵ *Ibid.*, 169.

of error.²⁶ The *Toronto Statement* also addressed some of the enforcement issues by urging the funders to make data management plans a necessary part of grant applications and made them subject to peer-review process.²⁷

2.1.4 EC recommendation C/2008/1329

The *EC recommendation C/2008/1329* of 10 April 2008 provided Member States with some important policy guidelines for the development/ updating of their national guidelines and frameworks.²⁸ One of the specific recommendations for Member States in this regard was to “[p]romote the broad dissemination of knowledge created with public funds, by taking steps to encourage open access to research results, while enabling, where appropriate, the related intellectual property to be protected”.²⁹

The *EC recommendation* also provided a *Code of Practice* for universities and other public research organisations, in order to improve the way public research organisations manage intellectual property and knowledge transfer. This *Code of Practice* for universities and other public research organisations, specifically suggests the development of a publication/dissemination policy that can promote broad dissemination of research and development results through methods like open access publication.³⁰ The *EC recommendation* also suggested the delays in such processes to be kept to a minimum.³¹ The *EC recommendation* also promoted open access to research data, subject to restrictions linked to commercial exploitation.³²

2.1.5 OECD guidelines for access to research data from public funding

Organisation for Economic Cooperation and Development (OECD) is an important inter-governmental organisation with 34 member countries, including many of the European

²⁶ Ibid., 170.

²⁷ Ibid., 169.

²⁸ *EC recommendation C/2008/1329* of 10 April 2008, <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32008H0416:EN:NOT> (accessed 11 December 2012).

²⁹ See Recommendation 4 of the *EC recommendation C/2008/1329*, Ibid.

³⁰ See Principle 7 in Annex 1 of the *EC recommendation C/2008/1329* of 10 April 2008, Ibid.

³¹ Ibid.

³² See Principle 16 in Annex II of the *EC recommendation C/2008/1329* of 10 April 2008, Ibid.

countries.³³ The *OECD Principles and Guidelines for Access to Research Data from Public Funding*, published in the year 2007, strongly recommends wider and open sharing of research data from public funding.³⁴

The specific aims and objectives of the *OECD Principles and Guidelines* include promoting “[a] culture of openness and sharing of research data among the public research communities within member countries and beyond” as well as stimulating “the exchange of good practices in data access and sharing”.³⁵

One of the key principles promoted by the *OECD Principles and Guidelines* is ‘openness’ and it insists “[a]ccess on equal terms for the international research community at the lowest possible cost, preferably at no more than the marginal cost of dissemination”.³⁶ The *OECD Principles and Guidelines* also recommends that “[o]pen access to research data from public funding should be easy, timely, user-friendly and preferably Internet-based”.³⁷

2.1.6 Implementation of the data policies and some of the important challenges

Review of research activities in genomic research show that the academic research community has mostly adhered to most of the early release obligations set out in data policies discussed above. This sub-section illustrates some of such initiatives.

One of the good examples in this regard is the *Fungal Genome Initiative* (FGS).³⁸ This initiative, led by the Broad Institute of Harvard and MIT, aims to provide the sequences of key organisms and their related species across the fungal kingdom.³⁹ This is with the aim of providing a strong foundation for works in medicine, agriculture, and industry through comparative studies.⁴⁰ *Fungal Genome Initiative* makes available the sequence data for all

³³ <http://www.oecd.org/about/membersandpartners/> (accessed 12 December 2012).

³⁴ See, generally, OECD, *OECD Principles and Guidelines for Access to Research Data from Public Funding*, <http://www.oecd.org/science/scienceandtechnologypolicy/38500813.pdf> (accessed 12 December 2012).

³⁵ *Ibid.*, 11.

³⁶ *Ibid.*, 15.

³⁷ *Ibid.*

³⁸ See <http://www.broadinstitute.org/scientific-community/science/projects/fungal-genome-initiative/fungal-genome-initiative> (accessed 10 December 2012).

³⁹ See <http://broadinstitute.org/science/projects/fungal-genome-initiative/frequently-asked-questions> (accessed 10 December 2012).

⁴⁰ *Ibid.*

FGI genomes in advance of assembly, in accordance with the NHGRI policy on rapid data release, by regular deposition of traces at the NIBI trace repository.⁴¹

Yet another good example is the *Marine Microbiology Initiative*, from the Moore Foundation.⁴² This initiative aims to “[u]ncover the principles that govern the interactions among microbes (who interacts with whom, how, when, where and the consequences thereof) and that influence the microbial mediated nutrient flow in the marine environment (who consumes and excretes what, where, how much, when and the consequences thereof)”.⁴³ The initiative insists on the release of DNA sequencing, assembly data, and annotations, along with associated metadata, within fifteen days, to the Community Cyber infrastructure for Advanced Marine Microbial Research and Analysis (CAMERA), which manages data release to collaborating laboratories and the general public.⁴⁴ The research groups submitting samples to this initiative are given exclusive access to the resulting sequence and annotation for an embargo period of six months and after this period, the data are made publicly available through CAMERA and NCBI.

The compliance with early data release norms for genomic sequences is found to be very high at the government level also. A good example in this regard is the pathogen genomics initiatives under the aegis of the National Institute for Allergy and Infectious Diseases (NIAID).⁴⁵ NIAID requires all raw genomic data and next generation sequencing data to be submitted as rapidly as possible either at the Trace Archive or, as appropriate, to the short read archive at NCBI/HLM/NIH.⁴⁶ Similarly, the recently started Genomic Encyclopedia of Bacteria and Archaea (GEBA) at the U.S. Department of Energy, also intends to release all genome sequence data for the community through DOE’s Joint Genome Institute (JGI) and GenBank.⁴⁷

Yet another important initiative at the governmental level is the International Human Microbiome Consortium (IHMC).⁴⁸ The primary objective of IHMC is “[t]o work under a common set of principles and policies to study and understand the role of the human microbiome in the maintenance of health and causation of disease and to use that knowledge

⁴¹ *FGI White Paper* (2003), 8, http://www.broadinstitute.org/fungi/fgi_02_whitepaper_2003.pdf

⁴² See <http://www.moore.org/marine-micro.aspx> (accessed 10 December 2012).

⁴³ *Ibid.*

⁴⁴ See <http://camera.calit2.net/> (accessed 10 December 2012).

⁴⁵ See <http://www.niaid.nih.gov/topics/pathogenomics/pages/relatedinitiatives.aspx> (accessed 10 December 2012).

⁴⁶ See <http://www.niaid.nih.gov/LabsAndResources/resources/dmid/gsc/Pages/data.aspx.html>

⁴⁷ See <http://www.jgi.doe.gov/programs/GEBA/> (accessed 10 December 2012).

⁴⁸ See <http://www.human-microbiome.org/> (accessed 11 December 2012).

to improve the ability to prevent and treat disease”.⁴⁹ For this goal, the Consortium focus on “[g]enerating a shared comprehensive data resource that will enable investigators to characterize the relationship between the composition of the human microbiome (or of parts of the human microbiome) and human health and disease”.⁵⁰

While the above examples provide a rosy picture with regard to liberal pre competitive access to data, the discussion should not lead one to think that implementation of open access initiatives are easy. One of the major challenges on the formation of an open knowledge environment as envisaged by most of the data policies is obtaining scientific information and data on terms consistent with the desired open access policies.

While many different databases available today provides open access and unrestricted use of data from direct providers and third-party database sources, majority of the data contributors are public funded entities (or individual researchers who are funded by public funded institutions) or non-profit entities. But the increasing pressures to commoditize databases and the evolution of database legislation in the EU may change this position.⁵¹ Yet another important challenge to be considered is the enforcement of contractually imposed attribution duties for all data, in all jurisdictions.

2.2 Marine data and UNCLOS

The United Nations convention on the Law of the Sea (UNCLOS) is considered to be the legal framework for all the activities conducted in the oceans. But the Convention itself does not make any direct reference to microbial organisms or genetic resources, as the available scientific knowledge on genetics was limited during the time of drafting of the Convention (between 1973 and 1982). However, some of the important obligations under the Convention relating to marine scientific research are applicable to modern large-scale database management also.

Notwithstanding the fact that genetic resources are not mentioned by the UNCLOS and given that customary international law recognizes the sovereignty of a state over its natural

⁴⁹ Ibid.

⁵⁰ Ibid.

⁵¹ See, for example, Schofield et al, *Nature* (2009) and Bryn Nelson, *Nature* (2009)

resources (first affirmed by the United Nations General Assembly in 1977⁵² and reaffirmed in the Preamble of the Convention on Biological Diversity) they are subject to the national sovereignty of the state where they are found. Therefore, as far as jurisdiction is concerned, the marine genetic resources that are within national jurisdiction are under the jurisdiction of the coastal state. Within the territorial sea (12 nautical miles from the baseline), the coastal state exercise sovereignty over natural resources.⁵³ In the exclusive economic zone (200 nautical miles from the baseline), the coastal state has the “*sovereign rights for the purpose of exploring and exploiting, conserving and managing the natural resources*”.⁵⁴ In the continental shelf (“*the seabed and subsoil of the submarine areas that extend beyond its territorial sea throughout the natural prolongation of its land territory to the outer edge of the continental margin*”), the coastal state has “*sovereign rights for the purpose of exploring it and exploiting its natural resources*”.⁵⁵

In areas beyond national jurisdiction (the high sea and the Area), the only natural resources that are managed by the UNCLOS are minerals of the seabed, subject to the legal regime of *common heritage of mankind*. This gave rise to what was then called the deepest of the ironies⁵⁶: minerals that were seen as the next “*blue gold*” to be exploited were regulated, but their exploitation was discovered to be technologically not feasible, while marine genetic resources that were not regulated were becoming the new promising “*blue gold*” to be exploited more feasibly. Marine genetic resources beyond national jurisdiction has also become the core of a highly political debate to determine which the applicable legal regime is and how they can be managed, given their crucial importance for future research in the field of upstream research and health and food applications, from which the whole humanity could benefit.

While the UNCLOS does not directly deal with pre-competitive access to marine research data, many of the provisions relating to marine scientific research (Part XIII, Articles 238-265) are relevant for pre-competitive access to marine research data. It may be noted here that the term ‘marine scientific research’ is not defined in the UNCLOS. But one section has defined the various aspects it has to comply with, notwithstanding the jurisdictional zone

⁵² General Assembly resolution 1803 (XVII) of 14 December 1962, "Permanent sovereignty over natural resources"

⁵³ Article 2 of the UNCLOS.

⁵⁴ Article 56 of the UNCLOS.

⁵⁵ Article 77.1 and Article 77.4 of the UNCLOS.

⁵⁶ Lyle Glowka, ‘The Deepest of Ironies: Genetic Resources, Marine Scientific Research, and the Area’ (1996) 12 *Ocean Yearbook* 154.

where the research is undertaken.⁵⁷ UNCLOS prescribes on the states and the competent international organizations a general duty to promote international cooperation in marine scientific research for peaceful purposes, in accordance with the principle of respect for sovereignty and jurisdiction and on the basis of mutual benefit.⁵⁸

UNCLOS also obliges the states and the competent international organizations to disseminate through appropriate channels the information on proposed projects as well as knowledge resulting from marine scientific research and “*promote the flow of scientific data and information and the transfer of knowledge resulting from marine scientific research, especially to developing states*”.⁵⁹ Finally, one may also note that, as far as marine scientific research in the Area (the seabed beyond national jurisdiction) is concerned, UNCLOS requires that it “[s]hall be carried out for the benefit of mankind as a whole”.⁶⁰ All such cooperation duties and obligations that promote the flow of scientific data and information make one consider UNCLOS as an important legal instrument for promoting more open access to large-scale microbial genomic databases.

2.3 Copyright and database rights

Two important branches of intellectual property law which have direct implications on access to data are copyright law and database rights.

Copyright is a branch of intellectual property law that protects original works of authorship in literary, dramatic, musical, and artistic works.⁶¹ Copyright is conferred automatically on creation of such works without any formal registration requirements and the threshold of creativity required for achieving copyright protection is very low. Copyright provides the authors (or the subsequent copyright owners, as the case may be) a bundle of rights which includes the right to reproduce the work, distribute copies of the work, and make derivatives based on the work.

⁵⁷ See Article 240 of the UNCLOS. This includes the need for research to be exclusively for peaceful purposes, with appropriate methods and means compatible with the Convention, in compliance with the relevant regulations for the protection of the environment and without creating unjustifiable interference with other legitimate uses of the sea.

⁵⁸ See Art. 242

⁵⁹ See Art. 244

⁶⁰ See Art. 143

⁶¹ <http://www.copyright.gov/help/faq/faq-general.html> (accessed 10 December 2012).

But copyright does not protect facts or ideas *per se* and it only protects their expression.⁶² Hence ‘raw data’ may not get any protection under copyright law. But their particular selection and arrangement may attain protection under copyright law.

Copyright law is not yet harmonised at the international level and the scope as well as enforcement of copyright is still determined by national legislation. But certain minimum standards with regard to copyright protection were laid down by the Berne Convention⁶³ and the Trade Related Aspects of Intellectual Property Rights (TRIPS) Agreement⁶⁴, which was part of the Agreement establishing the World Trade Organisation.⁶⁵ The WIPO Copyright Treaty (WCT), which was established under the aegis of the World Intellectual Property Organisation, has also laid down certain minimum standards in the area of copyright law, particularly with regard to protection of technological protection measures applied by copyright holders.⁶⁶ The copyright law is not completely harmonised within the European Union and one of the major initiatives at the community level in this regard was the Information Society Directive of 2001.⁶⁷ However, as mentioned earlier, the scope of protection and enforcement of copyright is still decided at the national level.

But within the European Union, databases may also enjoy an additional *sui generis* database right, by virtue of the Database Directive of 1996.⁶⁸ The directive defines ‘database’ as “[a] collection of independent works, data or other materials arranged in a systematic or methodical way and individually accessible by electronic or other means.”⁶⁹ The directive mandated the member states to provide a *sui generis* right “[f]or the maker of a database which shows that there has been qualitatively and/or quantitatively a substantial investment in either the obtaining, verification or presentation of the contents to prevent extraction and/or re-utilization of the whole or of a substantial part, evaluated qualitatively and/or

⁶² Ibid.

⁶³ See Berne Convention for the Protection of Literary and Artistic Works, http://www.wipo.int/treaties/en/ip/berne/trtdocs_wo001.html (accessed 10 December 2012).

⁶⁴ See Trade Related Aspects of Intellectual Property Rights (TRIPS) Agreement, http://www.wto.org/english/docs_e/legal_e/27-trips_01_e.htm (accessed 10 December 2012).

⁶⁵ The TRIPS Agreement is Annex 1C of the Marrakesh Agreement, which established the World Trade Organization.

⁶⁶ See WIPO Copyright Treaty, http://www.wipo.int/treaties/en/ip/wct/trtdocs_wo033.html (accessed 10 December 2012).

⁶⁷ Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society, <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32001L0029:EN:HTML> (accessed 10 December 2012).

⁶⁸ Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases (hereafter referred to as ‘Database Directive’), <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31996L0009:EN:HTML>

⁶⁹ See Art. 1 (2) of the Database Directive.

quantitatively, of the contents of that database".⁷⁰ In other words, the *sui generis* database right under the database directive arises when there has been substantial investment in obtaining, verifying or preserving the contents of a database.⁷¹

The Database Directive defines 'extraction' as "*the permanent or temporary transfer of all or a substantial part of the contents of a database to another medium by any means or in any form*" and 're-utilization' as "*any form of making available to the public all or a substantial part of the contents of a database by the distribution of copies, by renting, by on-line or other forms of transmission. The first sale of a copy of a database within the Community by the right holder or with his consent shall exhaust the right to control resale of that copy within the Community*".

The *sui generis* right provided by directive is to apply irrespective of the eligibility of the concerned databases for protection by copyright or any other rights.⁷² The new right is also applicable irrespective of the eligibility of the contents of that database for protection by copyright or by other rights and it is also without any prejudice as to the rights existing in respect of such contents.⁷³ The term of protection stipulated in the Directive is fifteen years from the date of completion of the making of the database.⁷⁴ Any "*substantial changes*" (evaluated qualitatively or quantitatively) to the database, "*including any substantial change resulting from the accumulation of successive additions, deletions or alterations*" may qualify as "*substantial new investment*", thereby giving the resulting database new term of protection.⁷⁵

While the Directive attempted to provide the broadest possible protection for the makers of databases, the interpretations given by the European Court of Justice with regard to the scope of database rights have considerably narrowed down the scope of protection. For example, in *British Horseracing Board v. William Hill Organisation Ltd.*, the ECJ clarified that creation of data do not qualify as 'obtaining' within the terms of the Database directive and that 'obtaining' involved seeking out existing independent materials and collecting them in the

⁷⁰ See Art. 7 (1) of the Database Directive.

⁷¹ Hector MacQueen, Charlotte Waelde, Graeme Laurie, and Abbe Brown, *Contemporary Intellectual Property: Law and Policy* (Oxford: Oxford University Press, 2010), 215.

⁷² See Art. 7(3) of the Database Directive.

⁷³ See Art. 7(3) of the Database Directive.

⁷⁴ See Art. 10 (1) of the Database Directive.

⁷⁵ See Art. 10 (3) of the Database Directive

database.⁷⁶ The Court specifically noted that “[t]he purpose of the protection by the *sui generis* right provided for by the directive is to promote the establishment of storage and processing systems for existing information and not the creation of materials capable of being collected subsequently in a database”. Nevertheless, practitioners observe that the *sui generis* right under the Directive provides strong protection for the database owners where the owner is the sole source of information contained in the database and in cases where such information could not have been collected independently in an economically viable way by third parties.⁷⁷

2.4 Data License Agreements

Data license agreements are used for imposing conditions regarding the access and use of the concerned data. While some use such agreements for preventing the passing of data from the first user to any third parties, some may also use such agreements for imposing additional restrictions that could impose considerable control over downstream innovations.⁷⁸ The data policies discussed in Section 2.1 have considerably influenced the formation of liberal data license agreements and many data license agreements today waive all rights over one's data while releasing them for the public.

Data license agreements can also be an important tool of legal certainty for data producers (and database providers), as legally valid waiver of rights (equivalent to a voluntary dedication to the public) from data sources or common-use licenses applicable for all incoming datasets are very important for assuring legal certainty with regard to downstream uses of the data.⁷⁹ Such efforts would also ensure other necessary practical aspects like ensuring the interoperability of the data submitted and thereby the ultimate goal of more open and liberal access to data.⁸⁰

⁷⁶ See *British Horseracing Board v. William Hill Organisation Ltd.*, Case: C-203/02, paras 29-31, <http://curia.europa.eu/juris/liste.jsf?language=en&num=C-203/02#> (accessed 10 December 2012).

⁷⁷ MacQueen, Hector, Waelde, Charlotte, and Laurie, Graeme (2010), *Contemporary Intellectual Property: Law and Policy* (Oxford: Oxford), 215-216 and 951.

⁷⁸ MacQueen, Hector, Waelde, Charlotte, and Laurie, Graeme (2010), *Contemporary Intellectual Property: Law and Policy* (Oxford: Oxford), 219-222 and 951.

⁷⁹ GEOSS, *Legal Options*, supra note __ §§2.6, 2.7

⁸⁰ Cf. id. §2.7.

2.5 Practical implications

The most important practical implication that one may notice from the discussions in this section, particularly in the context of the objective of this Deliverable to develop model agreements for pre-competitive access to microbial research data generated in marine bioprospecting missions, is the necessity to take into consideration the strong desire of the research community for free and liberal access to research data. This is very much evident from the various global and regional initiatives discussed in Sec. 2.1 and also the highly influential role played by such policies in global research exchanges today. The fact that the global research community has by and large adhered to the balancing approaches advocated by such policies further testifies the importance of complying with those principles while drafting the model agreements envisaged under the Deliverable. The model agreement developed as part of this Deliverable (provided in Sec. 4 of this Report) and also the data related provisions in the Model ABS Agreement (developed under Deliverable 8.1) are drafted by giving due recognition to the balancing approach advocated by such initiatives, particularly the *Fort Lauderdale Principles*.

Section 3. Two case studies

This section provides two interesting case studies that can positively influence the development of model agreements for pre-competitive access to large-scale microbial genomic research databases.

3.1 Case study 1 – Database of British Oceanographic Data Center

British Oceanographic Center (BODC) is a national facility in the United Kingdom for collecting and distributing data concerning marine environment. BODC maintains biological, chemical, physical and geophysical data and the databases contain measurements of nearly

22,000 different variables.⁸¹ The key aspect of the data policy of BODC is that it encourages the use of their data by all segments of the public, including academia and industry.⁸²

BODC make their data available under a license agreement.⁸³ If the data concerned was originated from the Natural Environmental Research Council (NERC), the data will be distributed according to the NERC Data Policy (See Annex 1).⁸⁴ The NERC Data policy generally suggests free distribution of data, except under a few circumstances wherein third parties hold rights over data.⁸⁵

If the original data was received from non-NERC organisations, the broader policy adopted is similar.⁸⁶ But the exact conditions are explained to the users after a request for the delivery of data is made.⁸⁷ By combining this hybrid approach, BODC has been able to bring substantial quantity of data to the public domain for furthering research and innovation.

3.2 Case study 2: BONUS EEIG Data Policy

The BONUS European Economic Interest Grouping (EEIG) is a program, jointly funded by different national funding organisations and the European Commission, with the specific goal of supporting the sustainable development of Baltic Sea and better management of Baltic Sea environmental issues.⁸⁸ This program funds research in many disciplines and generate a wide range of data from basic oceanographic field data to data from questionnaires, media & archive surveys and interviews.⁸⁹

The data policy of BONUS EEIG data policy is explicitly based on the principles advocated by the Organisation for Economic Cooperation and Development (OECD) through its report entitled *Promoting access to public research data for scientific, economic and social*

⁸¹ See http://www.bodc.ac.uk/about/what_is_bodc/ (accessed 10 December 2012).

⁸² https://www.bodc.ac.uk/data/where_to_find_data/ (accessed 10 December 2012).

⁸³ Ibid.

⁸⁴ Ibid.

⁸⁵ Ibid.

⁸⁶ Ibid.

⁸⁷ Ibid.

⁸⁸ See http://www.bonusportal.org/about_bonus/frequently_asked_about_bonus#WhatisBONUSSEEIG (accessed 12 December 2012).

⁸⁹ See 'BONUS EEIG data policy General Principles', http://www.bonusportal.org/files/1446/2012_01_19_BONUS_data_policy_Final.pdf (accessed 12 December 2012).

development (discussed in Sec. 2.1.5 of this report) and also the EC directive on public access to environmental information.⁹⁰ The two core principles that guide the BONUS EEIG data policy are that “*the public funded basic precompetitive research data (1) are a public good, produced in the public interest and (2) should be openly available to the maximum extent possible*”.⁹¹ The data policy of BONUS EEIG is an integral part of the BONUS intellectual property rights (IPR) regime and it consists of three sets of rules- the data policy statement, Articles 24 to 32 of the Grant Agreement, and possible stipulations in the Consortium Agreement. The data policy is annexed to the Grant Agreement to make it legally binding.

The data policy mandates global sharing of the data generated in BONUS programme funded projects and specifically mentions that they shall be made available to the scientific community with as few restrictions as possible. The suggested methods in this regard include storing of the original data in common databases and publishing of the research results in open access, as far as possible.⁹² BONUS EEIG data policy also mandates that all the data collected within the projects funded by BONUS have to be accompanied with contextual information or documentation (metadata) that can give secondary users all necessary details on the origin or manipulation of the data, so as to prevent any misuse, misinterpretation or confusion.⁹³

The BONUS Secretariat is entrusted with the collection and maintenance of metadata information on all data collected within projects.⁹⁴ Submission of metadata to the BONUS Secretariat is an obligatory part of the projects’ annual scientific reporting and the metadata will be available for the public in the BONUS portal site immediately after the delivery.⁹⁵ The BONUS EEIG data policy also provides specific directions regarding the sharing of data among the researchers within the BONUS program and stipulates that the refusal to follow the recommendations of the steering committee in the event of a dispute regarding sharing of data

⁹⁰ Ibid.

⁹¹ Ibid.

⁹² See ‘Obligations of global sharing of the data generated in BONUS programme funded projects’, http://www.bonusportal.org/files/1446/2012_01_19_BONUS_data_policy_Final.pdf (accessed 12 December 2012).

⁹³ Ibid.

⁹⁴ Ibid.

⁹⁵ Ibid.

may even result in termination of funding for the on-going project and rejection of new applications from the concerned researcher.⁹⁶

3.3 Practical Implications

Both the case studies discussed in this section illustrate different pragmatic approaches taken by data producers and data providers, with regard to sharing of research data. Both the case studies also demonstrate that such approaches can balance the interests of different stakeholders involved in the scientific research process. They also illustrate that ‘open and liberal access to research data’ is not merely an idealistic approach, but also a practically and technically feasible approach. Case studies like this are highly useful resources for the drafting of any robust and efficient model agreements for pre-competitive access to microbial research data generated in marine bioprospecting missions. Hence these case studies have played a highly positive part in the evolution of the provisions in the model agreement developed as part of this Deliverable and also the data related provisions in the model ABS agreement (developed under Deliverable 8.1), which are provided in the next section of this Report.

Section 4. Model Agreement and incorporation of the core principles in the ABS Agreement

Based on the lessons from the various analyses reported in Sec. 2 and Sec. 3 of this report, a model agreement that fulfils all the necessary conditions outlined under Work Package 8.2 has been developed. Some specific provisions relating to data are also included in the model ABS Agreement developed under Work Package 8.1, based on those insights. This section introduces the model agreement and also the specific data related provisions included in the model ABS Agreement.

⁹⁶

See ‘Principles of data sharing within BONUS programme’, http://www.bonusportal.org/files/1446/2012_01_19_BONUS_data_policy_Final.pdf (accessed 12 December 2012).

The model agreement has also attempted to combine the best licensing approaches existing in different organisations today, including the Creative Commons Licensing models and the EBI licensing model, so as to provide clear, easy to understand, and legally sound provisions for pre-competitive access to microbial research data generated from marine bioprospecting missions. This license may be used directly in the websites or the databases of the concerned missions as a click-wrap agreement, so as to achieve the objectives outlined. The complete model agreement is given below:

.....

Data License Agreement (Ver. 3.0)

License

THE TERMS AND CONDITIONS DESCRIBED IN THIS DATA LICENSE AGREEMENT ("LICENSE") APPLY TO ALL THE DATA GENERATED DURING _____, ORGANISED BY _____ ("LICENSOR"), WHICH ARE ACCESSIBLE THROUGH THIS DATABASE.

DATA (AS DEFINED BELOW) ARE PROVIDED BY THE LICENSOR TO YOU UNDER THE TERMS AND CONDITIONS OF THIS LICENSE. THE DATABASE IS PROTECTED BY COPYRIGHT AND/OR OTHER APPLICABLE LAW. ANY USE OF THE DATA OTHER THAN AS AUTHORIZED UNDER THIS LICENSE IS PROHIBITED.

BY ACCESSING THE DATABASE, YOU ACCEPT AND AGREE TO BE BOUND BY THE TERMS OF THIS LICENSE. TO THE EXTENT THIS LICENSE MAY BE CONSIDERED TO BE A CONTRACT, THE LICENSOR GRANTS YOU THE RIGHTS CONTAINED HERE IN CONSIDERATION OF YOUR ACCEPTANCE OF SUCH TERMS AND CONDITIONS.

1. Definitions

"Data" means all the genomic sequence data arising from _____, organised by the Licensor and which is made available for public access in this database.

"Proprietary use of the data" means any use of the data wherein the results of the use, including products developed using the data, are

not available to the public free of cost and without any restrictions.

"Public domain use of the data" means any use of the data wherein the results of the use are in the public domain and the public is free to use the results free of cost and without any restrictions.

"You" means an individual or entity exercising rights under this License who has not previously violated the terms of this License with respect to the data, or who has received express permission from Licensor to exercise rights under this License despite a previous violation.

2. License Grant

You are hereby authorised to access and use the data provided in this database for any public domain use.

3. Attribution requirement

You must provide appropriate attribution for all subsequent uses of the data in any product or service, including publications arising from the use of the data. Appropriate attribution shall be in the form of a byline/credit/link to both the Licensor and the original scientists/data contributors in the concerned campaign.

4. Compliance with the Convention on Biological Diversity and the Nagoya Protocol

Any use of the data under this license should be in compliance with the provisions of the Convention on Biological Diversity and the Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization.

5. Restrictions

The license granted in Section 2 above is also expressly made subject to and limited by the following restrictions:

5.1 The license granted is only for public domain use of the data. For any proprietary use of the data, prior written approval from Licensor is mandatory.

5.2 If You use data services of Licensor to a level that prevents, or looks set to prevent, Licensor providing services to others, Licensor may discontinue service to You. Licensor will then contact You to

discuss Your needs and how (and if) these can be met. Suggested solutions may include You licensing for Your own in-house use third-party software used in the services of Licensor. With permission from You, Licensor may divulge Your identity to such third parties, and jointly propose possible solutions.

5.3 You shall keep intact all notices that refer to this license and to the disclaimer of warranties.

6. Representations, Warranties and Disclaimer

6.1 Licensor does not guarantee the accuracy of any data or databases nor the suitability of databases, software and services for any purpose.

6.2 Except to the extent required by applicable law, in no event will Licensor be liable to You on any legal theory for any special, incidental, consequential, punitive or exemplary damages arising out of this license or the use of the work, even if Licensor has been advised of the possibility of such damages.

7. Data regarding access and usage

7.1 For planning, scientific review, and license compliance monitoring purposes, Licensor may keep records of usage of the data by You.

7.2 Licensor may make information about the total volume of usage of particular software or databases to third parties organisations who supply the software or databases.

7.3 Logs of usage may also be maintained for the purposes of monitoring and improving services, and measuring the impact a particular service has on the resources of Licensor. These logs will be kept in confidence and not made available for other purposes or to third parties.

7.4 Licensor regards the nature of the work carried out by You as confidential and shall make every reasonable attempt to prevent breaches of that confidentiality, except in cases of violation of the terms of this license by You and the subsequent proceedings relating to the violation of the terms of License. However, Licensor does not accept responsibility for the consequences of any breach of the confidentiality of the database of Licensor by third parties who by any means gain illegitimate access to information from Licensor.

8. Miscellaneous

8.1 Licensor itself places no intellectual property restrictions on any proprietary use of the data available via its services. However, some of the original data may be subject to patent, copyright, or other intellectual property rights claimed by third parties. It is Your responsibility to ensure that Your exploitation of the data does not infringe the rights of such third parties. Licensor can neither comment on the validity of their claims nor grant unrestricted permission regarding the use of the data.

8.2 Licensor will make all reasonable efforts to maintain continuity of service and provide adequate warning of any changes or discontinuities. However, Licensor accepts no responsibility for the consequences of any temporary or permanent discontinuity in service.

8.3 These Terms of Use reflect the enduring principles under which data are offered by Licensor. Licensor will not, without good reason, alter them, and where alterations are inevitable, Licensor will attempt to give reasonable notice of such alterations by publishing them in the website of the Licensor. However, Licensor do reserve the right to alter these terms at any time in order to ensure their appropriateness to the mission of the OSD project.

8.4 If any provision of this license is invalid or unenforceable under applicable law, it shall not affect the validity or enforceability of the remainder of the terms of this license, and without further action by the parties to this agreement, such provision shall be reformed to the minimum extent necessary to make such provision valid and enforceable.

8.5 No term or provision of this license shall be deemed waived and no breach consented to unless such waiver or consent shall be in writing and signed by the party to be charged with such waiver or consent.

8.6 You are strongly encouraged to contact the original scientists or data providers responsible for data made available by the Licensor. Where appropriate, researchers are also encouraged to consider collaboration and/or co-authorship with original investigators.

9. Termination

9.1 This License and the rights granted hereunder will terminate automatically upon any breach by You of the terms of this License.

9.2 Licensor may publicly publish the details of the violation of the terms of this license by You and also the details of the termination of License, in cases of violation of the terms of this License by You.

9.3 All disputes relating to this agreement shall be subject to the jurisdiction of _____.

This License constitutes the entire agreement between Licensor and You with respect to the data licensed here. There are no understandings, agreements or representations with respect to the data not specified here. Licensor shall not be bound by any additional provisions that may appear in any communication from You.

.....

The core principles of the model agreement are incorporated in the model ABS Agreement, developed under Work Package 8.1, to provide more legal certainty and also assurance to the provider countries that any data generated from the accessed genetic resources will be used for the wider benefit of the global scientific research community.

As mentioned in the preamble of the model ABS Agreement, one of the key objectives of the Agreement is to help develop innovative bioinformatics approaches to make the data accessible for the research and development community and for the public at large.⁹⁷

The ABS Agreements specifically seeks permission from provider countries for dissemination through online and offline media.⁹⁸ The possible forums suggested in the Agreement for online dissemination include the MicroB3 Portal (www.microb3.eu) and existing open data bases and information networks such as the Global Biodiversity Information Facility (GBIF), the Consortium for the Barcode of Life (CBOL), SeaDateNet, Pangaea and the International Nucleotide Sequence Databases (INSDC).⁹⁹

⁹⁷ See Preamble of the ABS Agreement.

⁹⁸ See Art. 6.1 of the Model ABS Agreement.

⁹⁹ Ibid.

The ABS Agreement also specifically mentions that the release of data resulting from research and development on the accessed genetic resources through online media, print media or delivery upon request will be organized such that users are bound not to use the specific data items directly linked to the accessed genetic resources and taken from the portals for proprietary purposes, unless they have obtained prior informed consent of the Provider State.¹⁰⁰

In order to further promote the culture of sharing of data within the research community, the ABS Agreement also specifically mentions the following in Article 6.3 of the Agreement: *“The Recipient may make data available also through the MicroB3 Information System. The users of such data shall be required to provide to the MicroB3 Information System a copy of the results of their research using the data accessed from the MicroB3 Information System, in such form and format as MicroB3 Information System will reasonably require in order to promote the objectives of research and development for the public domain.”* This provision may help considerably in the evolution of robust public domain, by promoting access to research results generated using the data provided through the MicroB3 Information System.

Section 5. Conclusion and Future Directions

As envisaged under the project proposal, Deliverable 8.2 has been able to map the model IP agreements for pre-competitive access to microbial research data generated in marine bioprospecting missions and also to propose legally valid provisions concerning data for the proposed model ABS Agreement developed under Work Package 8.1. The next two steps in the project are the development of the data policy guidelines and the fine tuning of the conditions in the model ABS Agreement with the help of the science community, so as to move towards the implementation of the model agreements for pre-competitive access to microbial genomic research databases. The latter will be done by conducting semi-structured interviews amongst the main business and science stakeholders. Such an effort would help to map the terms and conditions of the actual exchange practices. This would also help to identify the costs and benefits of the various model contracts for the stakeholders and society at large. A multi-stakeholder expert workshop will also be organized later to validate the

¹⁰⁰ See Art. 6.3 of the model ABS Agreement.

preliminary report of the survey and to build consensus on the contents of the proposed model agreements.

Annex 1

NERC Data Policy Statement

NERC has a policy on data in order to:

- a. Ensure the continuing availability of environmental data of long-term value for research, teaching, and for wider exploitation for the public good, by individuals, government, business and other organisations.
- b. Support the integrity, transparency and openness of the research it supports.
- c. Help in the formal publication of data sets, as well as enabling the tracking of their usage to be tracked through citation and data licences.
- d. Meet relevant legislation and government guidance on the management and distribution of environmental information.

NERC defines environmental data as individual items or records (both digital and analogue) usually obtained by measurement, observation or modelling of the natural world and the impact of humans upon it. This includes data generated through complex systems, such as information retrieval algorithms, data assimilation techniques and the application of models.

This policy covers environmental data acquired, assembled or created through research, survey and monitoring activities that are either fully or partially funded by NERC. It also applies to environmental data managed by NERC where NERC was not the original funder. This policy does not cover NERC's information products*.

This policy will be reviewed at regular intervals to ensure it keeps pace with scientific requirements and data management best practice.

Key principles

The environmental data produced by the activities funded by NERC are considered a public good and they will be made openly available for others to use. NERC is committed to supporting long-term environmental data management to enable continuing access to these data.

NERC will supply the environmental data it holds for free, apart from a few special cases as detailed in the policy.

NERC requires that all environmental data of long-term value generated through NERC-funded activities must be submitted to NERC for long-term management and dissemination.

Access to data

It is NERC's policy that:

1. All the environmental data held by the NERC Environmental Data Centres will normally be made openly available to any person or any organisation who requests them.
2. The only restrictions on access which we will apply are those supported by the exceptions on disclosure in the Environmental Information Regulations (2004). If it is proposed to restrict access to any data we will explain why.
3. To protect the research process NERC will allow those who undertake NERC-funded work a period to work exclusively on, and publish the results of, the data they have collected. This period will normally be a maximum of two years from the end of data collection.
4. All data held by the NERC Environmental Data Centres will be supplied for free except for large or complex requests where we may charge the cost of supply, or where third-party licence conditions either prevent such free supply, or require us to make specific charges.
5. All environmental data made available by the NERC Environmental Data Centres will be accompanied by a data licence. Data originally provided to NERC by a third-party may have their own access and licence conditions which restrict how or when we can make data available to others, in which case our data licence conditions will reflect these.
6. All those who use data provided by NERC are required to acknowledge the source of the data.

NERC's Environmental Data Centres

Successful long-term data management requires both specialist data curation skills and an understanding of the science behind the data. NERC achieves this by supporting Environmental Data Centres and co-locating these within its research and collaborative centres to take advantage of the scientific expertise they possess.

It is NERC's policy that:

7. NERC will maintain Environmental Data Centres for the management and dissemination of environmental data of long-term value generated through NERC funding or deposited by third-parties.
8. The data centres will act impartially towards all data producers, regardless of whether they are based within or outside of NERC. The environmental data within the data centres will be open to all on the same basis.
9. Working with the environmental science community NERC will maintain criteria to identify environmental data of long-term value (a Data Value Checklist). These criteria will be used to inform all decisions that NERC makes on the acceptance and disposal of data by its data centres.
10. Information on all data held within the data centres will be made available through the NERC Data Discovery Service.

Data collection

NERC expects everyone that it funds to manage the data they produce in an effective manner for the lifetime of their project, and for these data to be made available for others to use with as few restrictions as possible, and in a timely manner.

It is NERC's policy that:

11. All applications for NERC funding must include an outline Data Management Plan, which must identify which of the data sets being produced are considered to be of long-term value, based on the criteria in NERC's Data Value Checklist. The funding application must also identify all resources needed to implement the Data Management Plan.
12. The outline data management plan will be evaluated as part of the standard NERC grant assessment process. All successful applications will be required to produce a detailed data management plan in conjunction with the appropriate NERC data centre.
13. All NERC-funded projects must work with the appropriate NERC data centre to implement the data management plan, ensuring that data of long-term value are submitted to the data centre in an agreed format and accompanied by all necessary metadata.
14. Data from NERC-funded activities are provided to the data centres on a non-exclusive basis without prejudice to any intellectual property rights. This is to enable NERC to manage and make openly available publicly funded research data.
15. Those funded by NERC who do not meet these requirements risk having award payments withheld or becoming ineligible for future funding from NERC.

Open access to data underpinning research publications

NERC considers that long-term, open access to the data that underpin research publications will help to ensure the integrity, transparency and robustness of the research record. Access to these data supports the fundamental scientific requirement of allowing others to confirm or challenge research results.

It is NERC's policy that:

16. All research publications arising from NERC funding must include a statement on how the supporting data and any other relevant research materials can be accessed.
17. For all research publications produced by NERC's own staff, the supporting data will be made available through the NERC data centres.

* In line with UK Government policy, NERC distinguishes between data and information products. NERC defines environmental data as individual items or records (both digital and analogue) usually obtained by measurement, observation or modelling of the natural world and human impacts upon it, including all necessary calibration and quality control. This includes data generated through complex systems, such as information retrieval algorithms, data assimilation techniques and the application of models. Whereas, information products are created by adding a level of intellectual input that refines or adds value to data through interpretation and/or combination with other data. Model codes are not covered by this policy.