

# Data Policy Recommendations for Biodiversity Data. EU BON Project Report

Willi Egloff<sup>‡</sup>, Donat Agosti<sup>§</sup>, David Patterson<sup>|</sup>, Anke Hoffmann<sup>¶</sup>, Daniel Mietchen<sup>#</sup>, Puneet Kishor<sup>‡</sup>, Lyubomir Penev<sup>□</sup>

<sup>‡</sup> Plazi, Bern, Switzerland

<sup>§</sup> [www.plazi.org](http://www.plazi.org), Bern, Switzerland

<sup>|</sup> University of Sydney, Sydney, Australia

<sup>¶</sup> Leibniz Institute for Research on Evolution and Biodiversity, Berlin, Germany

<sup>#</sup> EvoMRI Communications, Jena, Germany

<sup>□</sup> Pensoft Publishers & Institute for Biodiversity and Ecosystem Research, Bulgarian Academy of Sciences, Sofia, Bulgaria

Corresponding author: Lyubomir Penev ([penev@pensoft.net](mailto:penev@pensoft.net))

Reviewable

v1

Received: 12 Mar 2016 | Published: 12 Mar 2016

Citation: Egloff W, Agosti D, Patterson D, Hoffmann A, Mietchen D, Kishor P, Penev L (2016) Data Policy Recommendations for Biodiversity Data. EU BON Project Report. Research Ideas and Outcomes 2: e8458. doi: [10.3897/rio.2.e8458](https://doi.org/10.3897/rio.2.e8458)

## Abstract

There is a strong need for a comprehensive, coherent, and consistent data policy in Europe to increase interoperability of data and to make its reuse both easy and legal. Available single recommendations/guidelines on different topics need to be processed, structured, and unified. Within the context of the EU BON project, a team from the EU BON partners from Museum für Naturkunde Berlin, Plazi, and Pensoft has prepared this report to be used as a part of the Data Publishing Guidelines and Recommendations in the EU BON Biodiversity Portal. The document deals with the issues: (i) Mobilizing biodiversity data, (ii) Removing legal obstacles, (iii) Changing attitudes, (iv) Data policy recommendations and is addressed to legislators, researchers, research institutions, data aggregators, funders, and publishers.

## Keywords

open science, biodiversity, data policies, data mining

## Introduction

The EU BON project will build a substantial part of the **Group on Earth Observation's Biodiversity Observation Network (GEO BON)** to ensure sustainable governance of our biological resources. Regarding the development of the EU BON **Data Policy Recommendations (DPR)** (milestone MS972), there is an overlap between tasks 8.4 and 9.7. In task 8.4, the milestone MS841 '*Biodiversity data publishing legal framework report*' was submitted in May 2015 (Suppl. material 1), and in task 9.7, milestone MS971 '*Data sharing agreement*' was finalized in April 2014 (Suppl. material 2). In addition, the paper "*Open exchange of scientific knowledge and European copyright: The case of biodiversity information*" published in the open access journal ZooKeys (Egloff et al. 2014) and covers copyright issues across EU countries relating to biodiversity data publishing.

The main purpose of **Task 8.4** (Data Publishing, Data Citation, and Data Usage Strategy and Guidelines) is the implementation of a Strategy and Guidelines for peer-reviewed, open-access data publishing, citation and usage as an important incentive for authors to publish their data, thereby sharing them for subsequent re-use. A legal framework for data publishing and dissemination will be developed. Special emphasizes will be given to the development of peer-review strategies for research data.

The main purpose of **Task 9.7** (Data policies and Intellectual Property Rights) is to monitor national, European, and international policies, legal frameworks and provisions which may affect access to, management of, and, in particular, subsequent sharing and distribution of biodiversity data as far as is relevant to the EU BON project. Information about existing legal requirements and provisions affecting the EU BON data services will be made available to the project partners, and feedback will be sought on current practices with data handling and observing rights and legal obligations. As a major output, data policy recommendations will be formulated for EU BON.

## Progress towards objectives

We surveyed the copyright and usage licenses used by the potential suppliers of data to the EU BON portal listed in the Annex to milestone MS241 '*Specification for registry and metadata catalogue*' (Suppl. material 3). A summary of this survey, together with the data sharing agreement and the survey on Intellectual Property Rights (IPR) issues on biodiversity data in the European Union (Egloff et al. 2014 provides a good basis for the Data Policy Recommendations which constitute milestone MS972. For this, we have formulated a set of data policy recommendations based on:

1. EU BON data sharing agreement (MS971).
2. The paper of on the legal framework for biodiversity information in Europe (Egloff et al. 2014).
3. Analysis of IPR policies of the EU BON data suppliers (Annex of milestone MS241).

4. The Bouchout declaration principles, see the website (<http://bouchoutdeclaration.org>) or Fig. 1.

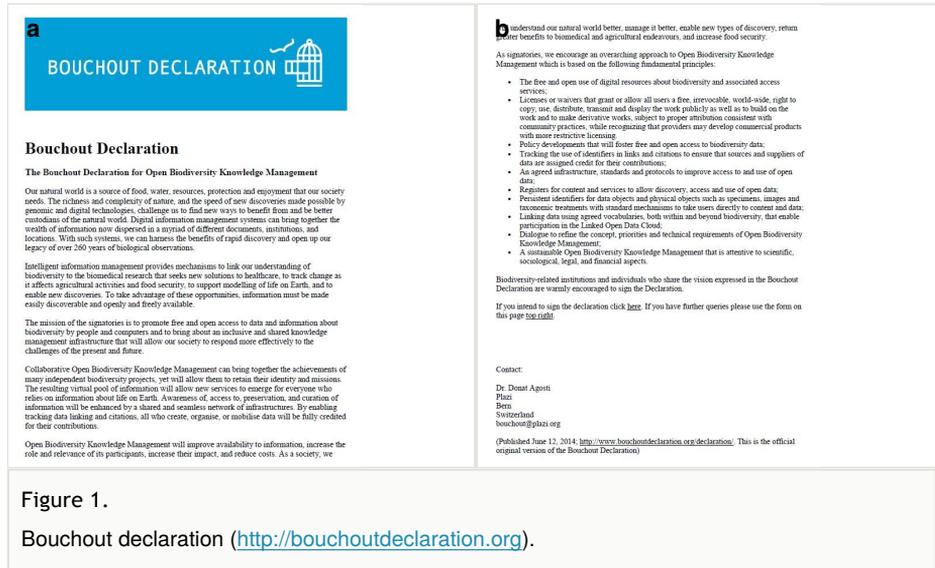


Figure 1.  
Bouchout declaration (<http://bouchoutdeclaration.org>).

5. A RECODE project [deliverable](#).

6. Pensoft's [Data Publishing Policies and Guidelines for Biodiversity Data](#) (Penev et al. 2011).

7. Official data policies statements and documents of major funders and research organizations (e.g. Horizon2020, National Science Foundation (NSF) and National Institutes of Health (NIH) of the USA, and others).

8. Other sources, cited within the document.

## Findings

### Background

Biodiversity data and information provide important knowledge for many biological, geological, and environmental research disciplines as well as for the development of policies relating to the natural environment and the management of natural resources. Digital information management systems can bring together the wealth of information and the legacy of over 260 years of biological observations now dispersed in a myriad of different documents, institutions, and locations. As the signatories of the [Bouchout Declaration for Open Biodiversity Knowledge Management](#) declare, “intelligent information management provides mechanisms to link our understanding of biodiversity to the biomedical research that seeks new solutions to healthcare, to track change as it affects

agricultural activities and food security, to support modeling of life on Earth, and to enable new discoveries. To take advantage of these opportunities, information must be made easily discoverable and openly and freely available.”

This standpoint is accepted by most public and scientific authorities. In 2007, the Organisation for Economic Co-operation and Development (OECD) published “[Principles and Guidelines for Access to Research Data from Public Funding](#)” that intend to facilitate the access to research data generated with public funding. In 2012, the European Union released a “[Commission Recommendation on access to and preservation of scientific information](#)”, which serves the same aim. Together with the European Federation of Academies of Sciences and Humanities, the European Commission published a “[Joint Declaration on Open Science for the 21st Century](#)”.

At present, the access and reuse of biodiversity data is hampered by an array of technical, economic, sociological, legal and other factors. Considerable quantities of biodiversity information are detained in legacy literature that is not accessible for technical reasons. Many compilers of biodiversity content act as if or claim that they hold intellectual property rights over their data and information. Open and free access to biodiversity data and information requires that we overcome these obstacles. It is to that end that EU BON has elaborated its data policy.

## **Mobilizing biodiversity data**

Biodiversity data can be mobilized from three different major sources of information:

- Raw data from observations/collections published via data aggregators and citizen science platforms
- Unlocking the printed legacy literature through conversion to a digital format, retrospective markup, and/or text and data mining
- Prospective markup of new publications

These three sources of biodiversity data each need appropriate policies and guidelines to incentivise data providers and custodian to publish the data. In spite of the diversity of specific national, institutional, domain-specific, and individual requirements and expectations regarding copyright and norms accepted and used across countries, we can formulate a few *strategic goals* that should be adopted and implemented for all three data sources (see Data Policy Recommendations section for more detail).

Strategic goals for biodiversity data mobilization and publication:

1. Promote the understanding that primary biodiversity data are facts and therefore NOT a subject of copyright; they belong to the public domain, independent of their source;
2. We should require explicit statements that clearly place biodiversity data in the public domain, by applying a standardized waiver for any eventual copyright or database protection right, for example [Creative Commons Zero \(CC0\)](#). Some

countries (cf. <https://github.com/unitedstates/licensing/issues/31>) may still need special licenses for data irrespective of its source.

3. To the maximum possible extent, we should render printed materials, PDFs, and other non-machine-actionable biodiversity data and narratives, into machine-readable and harvestable formats.

## Removing legal obstacles

No intellectual property rights apply to information or data. “Intellectual property rights” are a group of legal instruments that exist in many countries and are applied to precise immaterial goods in a precise context. In member countries of the EU, “intellectual property rights” refer mainly to copyright (conceived in relation to creative works of art and literature), neighboring rights (relating to performances, phonograms and broadcasts), patent rights (relating to inventions), industrial designs, trademarks and databases. The concept of intellectual property rights applies only to goods that are precisely defined: Where there is no law stipulating explicitly the protection of a specified class of immaterial items, no intellectual property rights exist.

Data and information in general, or biodiversity data in particular, are not protected immaterial goods. Consequently, there can be no intellectual property right on biodiversity data as such. A legal protection can only exist if the biodiversity data qualify as one of the protected immaterial goods. In practice, this can occur where collections of biodiversity data qualify as a “work” in the meaning of copyright or as a “database” in the meaning of EU database protection.

Copyright can be applied to works that are original, individual, new creations with respect to the form of the presentation. It does not cover ideas, procedures, systems nor content. Scientific data present facts in standardized forms that have been agreed by the respective scientific community. As they are not creative in form, scientific data in general as well as their metadata do not qualify as works. This is also valid for numerous biodiversity data presented as images because they present facts according to standardized, preconceived conventions.

On the other hand, copyright protection can apply to a collection of biodiversity data if it constitutes, by reason of the selection or arrangement of their contents, an intellectual creation with an individual character. The more systematic a collection of data is, and the more consistent with agreed standards and conventions, the less individual it is in the meaning of copyright, and the less likely copyright protection will apply. Consequently, collections of biodiversity data will be protected by copyright only in a very small minority of cases. Nevertheless, in these few cases, copyright may constitute a barrier to the free exchange of biodiversity data.

European copyright legislators are well aware of this impediment to data exchange. The [EU Directive 2001/29/EC](#) on the harmonisation of certain aspects of copyright and related rights in the information society addresses this challenge. It puts considerable weight on the importance of science by providing for exceptions and limitations to copyright. It grants

to the author the rights to decide who shall be allowed to reproduce his work (“reproduction right”) or who shall be allowed to communicate it to the public (“communication right”), but it provides also for several restrictions (“exceptions and limitations”) to intellectual property rights in the general interest. They refer, among others, to “educational and scientific purposes” (Recital 34) or to “the benefit of certain non-profit making establishments such as publicly accessible libraries and equivalent institutions, as well as archives” (Recital 40). However, these exceptions and limitations are only applicable if and when they are transformed into national law by individual member states of the EU, and in such cases, they apply only to that member state.

The EU Database protection is not part of copyright but is a *sui generis* (special case) right that applies whether copyright relating to the database exists or not. It applies only to databases which show “that there has been qualitatively and/or quantitatively a substantial investment in either the obtaining, verification or presentation of the contents” (art. 7, [Directive 96/9/EC](#)). As the European Court of Justice pointed out in several judgments, database protection concerns the creation of databases out of material that already exists, but does not deal with the creation of those data. The expression “investment in the obtaining of the contents” refers therefore to the resources used to find existing materials and collect them in the database, and not to the resources used to create materials. Databases, like scientific papers, collect data in categories agreed by the scientific community, apply domain-specific standards, and use standard protocols to make content accessible. That is, ‘the presentation of contents’ is rarely creative. Database rights only refer to the database as a whole, not to individual units of data. Database rights are violated by unauthorized use of the whole or substantial part of the database. Database rights do not prevent the use of individual data elements or minor parts of the data collection. The EU Database protection also provides for exceptions and limitations in the general interest, for example in the interest of scientific efforts. As in the case of copyright, these exceptions and limitations are only applicable when they are transformed into national law by individual member states of the EU, and in this case, they apply only to that member state.

As illustrated in a recently published review (Egloff et al. 2014), such transformations into national law have resulted in many differences among national practices. National provisions in Europe on copyright protection and the exceptions and limitations for research purposes differ not only in details but in substance. There is no consistency among national legislations despite Directive 2001/29/EC that aims to achieve harmonisation. Exceptions to the *sui-generis* database-protection are even more varied. Therefore, scientists who rely on data from different EU member states or who collaborate internationally need to be aware that different legal frameworks may apply to the data they use. In the Communication on “Copyright in the Knowledge Economy”, the EU Commission makes it clear that this situation is a major stumbling block to international scientific cooperation within the EU.

Copyright as well as database protection are part of “private law”, which is applied only on demand by the owner of the rights. Even if there is an intellectual property right with respect to a particular collection of biodiversity data, the owner is entitled to renounce their

claim to those rights. This principle of private law is the basis of the common-sense phrase: “Where there is no plaintiff, there is no judge”.

## Changing attitudes

The reluctance of researchers and publishers to distribute and exchange their data and information openly has economic, scientific, or sociological reasons (Thessen and Patterson 2011).

One factor that may change this reluctant attitude is to develop measures that ensure that all who create, organise or mobilise data are fully credited for their contributions (Patterson et al. 2014). This can be achieved by applying Universally Unique Identifiers (UUIs) to any element of data or information, and track the export and import of the content by, for example, using small plugins for browsers. That way, sources and suppliers of data can be assigned credit for their contributions by tracking the use of identifiers. Proper attribution is an established community norm for all scientific information, be it protected by any intellectual property rights or not. Therefore, the right of attribution does not require the recognition of any intellectual property right.

Biodiversity data and information should not be treated as commercial goods, but as a common resource for the whole human society. From this perspective, scientific publications need to be made openly available, as soon after publication and as freely as possible. Researchers should be able to communicate their results with minimum time delay and at minimum cost. Restrictions to open availability should only be applied if based on specific justifications, such as to protect security, endangered species, or to protect the privacy of individuals.

## Data Policy Recommendations

The main objective of EU BON is to build a substantial part of the Group on Earth Observation’s Biodiversity Observation Network (GEO BON). EU BON’s deliverables include a comprehensive „European Biodiversity Portal“ for all stakeholder communities, strategies for a global implementation of GEO BON and support of the Intergovernmental Platform on Biodiversity and Ecosystem Services (IPBES). In that perspective, EU BON recommends to all members, associated persons and institutions as well as to other stakeholders of biodiversity information to contribute to the following data policy:

### 1. Legislators

- The EU should revise the Directive 2001/29/EC by declaring that the provision of a copyright exception for scientific research is compulsory for all member states. The new regulations should not refer to commercial or non-commercial scientific research, as this distinction is neither useful nor applicable in practice. Nor should they refer to the place from where, nor the technical mode how, works are accessed, as such restrictions hamper the research process.

- The EU should revise the Directive 96/9/EC by declaring that the re-use of protected databases for scientific research is authorised by a compulsory exception to database rights.
- Member states of the EU or the EEA should introduce or, where it already exists, extend a copyright exception for the use of works for scientific research. This exception should not refer to commercial or non-commercial scientific research, as this distinction is neither useful nor applicable in practice. Nor should it refer to the place from where, nor the technical mode how, works are accessed, as such restrictions hamper the research process.
- Member states of the EU or the EEA should introduce or, where it already exists, extend an exception of database protection for the re-use of databases for scientific research.

## 2. Researchers

- Researchers should refrain from asserting intellectual property rights for biodiversity data and information collected and/or published by them. By default, all content referring to biodiversity information should be openly accessible.
- As far as material produced by researchers is protected by copyright or by database rights, the right owner should make these works or databases freely accessible and reusable by publishing them under a [CC-BY](#) or [CC0](#) .
- Publicly funded research institutions should refrain from asserting intellectual property rights for biodiversity data and information collected and/or published by them. By default, all content referring to biodiversity information should be openly accessible.
- Publicly funded institutions should encourage re-use of biodiversity data and information for research purposes with a requirement for attribution of the source, but should impose no other requirements on re-use.
- As far as material owned by publicly funded institutions is protected by copyright or by database rights, the institutions should dedicate these works or databases to the public domain by publishing them under [CC0](#) .

## 3. Data aggregators

- Encourage data suppliers and partner nodes to publish their data under [CC0](#). With [CC0](#), the data publisher waives any copyright over the data(set) and dedicates it to the public domain. Users can copy, use, modify and distribute the data without asking your permission. The data publisher cannot be held liable for any (mis)use of the data either. [CC0 is recommended for data and databases](#) and is used by hundreds of organizations. It is especially recommended for scientific data and thus encouraged by [Pensoft](#) (see, for example, the [policies](#) of the Research Ideas and Outcomes (RIO) journal); such an appeal has been published in Nature as early as in 2009 (Schofield et al. 2009). Although [CC0](#) doesn't legally require users of the data to cite the source, it does not take away community norms on the moral responsibility to give attribution, as is common in scientific research.

- Ensure that data are stored in a versioned and time-stamped manner.
- Provide data citation mechanisms (Starr et al. 2015) at the level of dataset and individual data records. Good examples are used by [Canadensys](#), [VertNet](#), Pensoft (Penev et al. 2011), and the [RDA Working Group on Data Citation](#) (WGDC).
- Develop mechanisms to identify and cite arbitrary views of data, from a single record to an entire data set in a precise, machine-actionable manner, that are stable across different technologies and technological changes.
- Allow to cite and retrieve that data as it existed at a certain point in time, whether the database is static or highly dynamic.
- Identify data sets by storing and assigning persistent identifiers (PIDs) to time-stamped queries that can be re-executed against the time-stamped data store.

#### 4. Funding Agencies

- Whenever possible, funders should support and require use of the most liberal data use and re-use licenses, particularly by putting data into the public domain through the [CC0](#).
- Develop policies to require funded researchers to make the data underpinning scientific publications available in machine readable formats in public repositories at the time of initial publication.
- Support and require enhancement to the maximum possible extent of machine-readability of both data and associated metadata.
- Ensure and require that data management plans (DMPs) to include clear statements and a work plan for archiving and sharing research data. The DMPs should include: descriptions of data to be produced in the proposed study, any data standards used, mechanisms for providing access to and sharing of data (including provisions for protection of privacy, confidentiality, security, intellectual property, or other rights), provisions for data reuse and redistribution, and plans for archiving and long-term preservation of the data. As the technical infrastructure available to research will continue to evolve, so we expect DMPs to evolve.
- Support and encourage the use of established public repositories and community based standards. Funders should encourage all funded researchers to make use of existing data standards relevant to their research community, such as standards for collecting and representing data and information describing the data set (i.e. metadata), as well as promote the interoperability of digital data in and across public repositories.
- Develop approaches and support technologies to ensure the discoverability of data to make them findable, accessible, and citable. Funders should support also the development of data discovery indexes to provide a mechanism to enhance discoverability and facilitate appropriate attribution to those responsible for the dataset and link the citations to associated publications.
- Explore the development of a data commons, a shared space for research output including data, software and a narrative that follows the FAIR principles of Find, Access, Interoperate and Reuse.

## 5. Publishers

- Allow authors to retain the copyright to their publications and to make it available under the terms of the [Creative Commons CC-Zero Waiver \(CC0\)](#), or if requested, of the [Creative Commons Attribution License \(CC BY 4.0\)](#) .
- Use the [Creative Commons CC-Zero Waiver](#) for data in any supplementary materials associated with an article.
- Other data publishing licenses may be allowed as exceptions (subject to approval by the editor on a case-by-case basis) and should be justified with a written statement from the author that will be published with the article.
- Authors should be required to share all data, code or protocols underlying the research reported in their articles. Exceptions could be permitted, but have to be justified in a written public statement accompanying the article.
- Datasets and software should be deposited and permanently archived in appropriate, trusted, general, or domain-specific repositories (please consult <http://service.re3data.org> and/or software repositories such as [GitHub](#), [SourceForge](#), [BitB ucket](#), [Bioinformatics.org](#), or equivalent). The associated persistent identifiers (e.g. DOI, or others) of the dataset(s) must be included in the data or software resources section of the article. Reference(s) to datasets and software should also be included in the reference list of the article with DOIs and other identification schemes (where available). Where no domain-specific data repository exists, authors should deposit their datasets in a general repository such as [ZENODO](#), [Dry ad](#), [Dataverse](#), or others.
- Small bodies of data may also be published as data files or packages supplementary to research articles, however, the authors should favour deposition in data repositories.
- Ensure availability of both data and narrative in harvestable machine-readable formats (for example JATS XML), including article and dataset metadata.
- Increase the proportion of machine-readable content within the narrative and data to the maximum extent possible at the current level of technology development.

## Other documents relating to Data Policy Recommendations

We here include reference to other documents that are relevant to the EU BON data policies and summaries of their content.

### **Biodiversity Data Publishing Legal Framework (EU BON Milestone MS841) (summary)**

On the basis of data policy principles, EU BON has agreed upon a data-sharing policy, which is binding for providers of data to the EU BON portal, EU BON partners, associated persons and institutions, and for all users of its portal. The report is part of the process to develop the legal framework on which this data sharing policy will be based.

The legal framework for data publishing and dissemination applicable to EU BON is realized in the form of the EU BON Data Sharing Agreement. By asking data providers to refrain from claiming intellectual property rights, it makes sure that no such rights are applied to data within the EU BON network. For data under national or international security restrictions or under time embargos, EU BON provides for a special category of “sensitive data”. Such data are kept separately from other data and are made available only upon special justification. Finally, EU BON does not assert any intellectual property rights for itself; it dedicates all collections of data that might qualify as works in the meaning of copyright to the public domain or publishes them under a Creative Commons (CC-BY) 4.0 license.

Content:

- Intellectual property rights on biodiversity data?
- Other legal aspects referring to biodiversity data
- The EU BON Data Sharing Agreement

**Data Sharing Agreement (EU BON Milestone MS971) (summary)**

The Data Sharing Agreement sets out the policy of EU BON on the sharing and use of data available in the EU BON portal. The document refers to EC policies ([Scientific data: open access to research results will boost Europe's innovation capacity](#)) and the GEOSS Data Sharing Principles and includes two paragraphs on intellectual property rights (2.3, 3.3).

Content:

- Background or guiding principles
- Obligations and guarantees for data providers
- Obligations and guarantees for EU BON
- Obligations and guarantees for users
- Contribution to GEO BON

**Paper on European Copyright Law with Respect to Biodiversity Data (summary)**

The paper by Egloff et al. (2014) largely covers the several issues in the EU BON taska 8.4 and 9.7, namely a survey of the biodiversity data publishing and copyright policies across EU countries.

Content:

- The request for open access
- Why copyright can hamper the exchange of biodiversity knowledge
- The European database protection
- The importance of data use agreements
- Examples of national regulations

## Specifications for Registry & Metadata Catalogue (EU BON Milestone MS241) (summary)

Several specific components and sub-networks have already been identified for integration in EU BON. In addition to GBIF, these include, for example, the broader DataONE network (including the Long Term Ecological Research program (LTER), the International Long Term Ecological Research networks (ILTER), the Knowledge Network for Biocomplexity (KNB), and the Dryad digital repository), the database of the EU-wide monitoring methods and systems of surveillance for species and habitats of community interest (DaEuMon) and the Drupal Ecological Information System (DEIMS). The registry lists network resources that are expected to be connected to the EU BON registry. For each entity, the access protocol, metadata standard and any accessor requirements are noted.

### The Bouchout Declaration Principles

The [Bouchout Declaration](#) (see also Fig. 1) was launched by the [pro-iBiosphere](#) FP7 project. The purpose of the Bouchout Declaration is to help make digital data about our global biodiversity openly available worldwide. It offers members of the biodiversity community a way to demonstrate their commitment to open science.

So far, the Bouchout Declaration has been signed by more than 90 organizations and more than 200 individuals.

### The RECODE Project Report

The EU [Project RECODE](#) ('Policy RECommendations for Open Access to Research Data in Europe') published:

- Deliverable (D5.1) on "[Policy guidelines for open access and data dissemination and preservation](#)" (Feb 2013), and
- Summary booklet on "[Policy recommendations for open access to research data](#)" (2014). These recommendations include among others also environment research data.

RECODE provides:

#### 1. Ten overarching recommendations:

- Develop aligned and comprehensive policies for open access to research data;
- Ensure appropriate funding for open access to research data;
- Develop policies and initiatives that offer researchers rewards for open access to high quality data;
- Identify key stakeholders and relevant networks and foster collaborative work for a sustainable ecosystem for open access to research data;
- Plan for the long-term, sustainable curation and preservation of open access data;

- Develop comprehensive and collaborative technical and infrastructure solutions that afford open access to and long-term preservation of high-quality research data;
- Develop technical and scientific quality standards for research data;
- Require the use of harmonized open licensing frameworks;
- Systematically address legal and ethical issues arising from open access to research data; and
- Support the transition to open research data through curriculum-development and training

2. Stakeholder-specific recommendations for funders, research institutions, data managers, publishers

3. Practical guides for these groups, including: (i) Preparing and implementing a policy; (ii) policy content; (iii) practical checklist for the specific group. At the end, they provide a long list of resources, including funder policies, EC policies for Open Access, publisher policies etc.

### **Pensoft Data Publishing Policies & Guidelines for Biodiversity Data**

This is an extensive document that provides the basis for the data publishing practices in Pensoft's journals and can be used by other publishers when appropriate (Penev et al. 2011).

#### Content:

- Data Publishing Policies
  - Data Publishing in a Nutshell
  - Why Publish Data
  - Data Publishing Licenses
  - How to Publish Data
  - Open Data Repositories
- How to Cite Data in Pensoft Journal Articles
- Guidelines for Authors
- Data Published within Supplementary Information Files
- What is a “Data Paper”
  - Data Papers Describing Primary Biodiversity Data
  - Generation of Data Paper manuscripts using the GBIF Integrated Publishing Toolkit (IPT)
  - Data Papers Describing Ecological and Environmental Data
  - Data Papers Describing Genome Data
  - Barcode Data Release Papers
  - Data Papers Describing Software Tools
- Guidelines for Reviewers of Data Papers
  - Quality of the Manuscript
  - Quality of the Data
  - Consistency between Manuscript and Data

## Project

EU BON: Building the European Biodiversity Observation Network, an European research project, financed by the 7th EU framework programme for research and development (FP7), grant agreement number 308454.

## References

- Egloff W, Patterson D, Agosti D, Hagedorn G (2014) Open exchange of scientific knowledge and European copyright: The case of biodiversity information. *ZooKeys* 414: 109-135. DOI: [10.3897/zookeys.414.7717](https://doi.org/10.3897/zookeys.414.7717)
- Patterson DJ, Egloff W, Agosti D, Eades D, Franz N, Hagedorn G, Rees JA, Remsen DP (2014) Scientific names of organisms: attribution, rights, and licensing. *BMC Research Notes* 7 (1): 79. DOI: [10.1186/1756-0500-7-79](https://doi.org/10.1186/1756-0500-7-79)
- Penev L, Mitchen D, Chavan V, Hagedorn G, Remsen D, Smith V, Shotton D (2011) Pensoft Data Publishing Policies and Guidelines for Biodiversity Data. Pensoft Publishers, [http://www.pensoft.net/J\\_FILES/Pensoft\\_Data\\_Publishing\\_Policies\\_and\\_Guidelines.pdf](http://www.pensoft.net/J_FILES/Pensoft_Data_Publishing_Policies_and_Guidelines.pdf). [http://www.pensoft.net/J\\_FILES/Pensoft\\_Data\\_Publishing\\_Policies\\_and\\_Guidelines.pdf](http://www.pensoft.net/J_FILES/Pensoft_Data_Publishing_Policies_and_Guidelines.pdf)
- Schofield P, Bubela T, Weaver T, Portilla L, Brown S, Hancock J, Einhorn D, Tocchini-Valentini G, de Angelis M, Rosenthal N (2009) Post-publication sharing of data and tools. *Nature* 461 (7261): 171-173. DOI: [10.1038/461171a](https://doi.org/10.1038/461171a)
- Starr J, Castro E, Crosas M, Dumontier M, Downs R, Duerr R, Haak L, Haendel M, Herman I, Hodson S, Hourclé J, Kratz JE, Lin J, Nielsen LH, Nurnberger A, Proell S, Rauber A, Sacchi S, Smith A, Taylor M, Clark T (2015) Achieving human and machine accessibility of cited data in scholarly publications. *PeerJ Computer Science* 1: e1. DOI: [10.7717/peerj-cs.1](https://doi.org/10.7717/peerj-cs.1)
- Thessen AE, Patterson DJ (2011) Data issues in the life sciences. *ZooKeys* 150: 15-51. DOI: [10.3897/zookeys.150.1766](https://doi.org/10.3897/zookeys.150.1766)

## Supplementary materials

### Suppl. material 1: MS841: Biodiversity data publishing legal framework report

**Authors:** W. Egloff, D. J. Patterson, D. Agosti, L. Penev

**Data type:** PDF document

**Brief description:** EU BON milestone report MS841

**Filename:** Milestone\_MS841\_Biodiversity Data Publishing\_Legal Framework.pdf - [Download file](#)  
(365.10 kb)

**Suppl. material 2: MS971: Data sharing agreement**

**Authors:** EU BON Partners

**Data type:** PDF document

**Brief description:** EU BON milestone report MS971

**Filename:** EU BON\_Data Sharing Agreement\_20140401 (1).pdf - [Download file](#) (110.88 kb)

**Suppl. material 3: MS241: Specification for registry and metadata catalogue**

**Authors:** EU BON Partners

**Data type:** PDF document

**Brief description:** EU BON milestone report MS241

**Filename:** Milestone\_MS241\_Specification for registry and metadata catalogue.pdf - [Download file](#) (739.02 kb)