



[PRO iBiosphere web site](#)



PRO iBIOSPHERE
WWW.PRO-IBIOSPHERE.EU



[Subscribe to this service](#)

NEWSLETTER

Issue 4

(September/December 2013)

In this Issue:

Commission launches pilot to open up publicly funded research data

16.12.2013

Today, 16/12/2013, the European Commission announced the launch of a new Pilot on Open Research Data in Horizon 2020, to ensure that valuable information produced by researchers in many EU-funded projects will be shared freely. Researchers in projects participating in the pilot are asked to make the underlying data needed... [more](#)

The pro-iBiosphere project highlighted at the 6th EU-AU Cooperation Forum on ICT in Addis Ababa

13.12.2013

The 6th Africa-EU Cooperation Forum on ICT took place on December 2-3 2013 at the African Union Conference Center in Addis Ababa, Ethiopia under the aegis of the European Commission and the African Union Commission. The event was organised by the EU EuroAfrica-P8 project on the occasion of the 50th anniversary... [more](#)

Working, not drinking, at the Taverna

10.12.2013

Data analysis is a complicated and time-consuming process. Like a craftsman, you require a set of tools that source, reformat, merge and analyse data. Using these tools manually in a workflow can take weeks. Then, when you finally get the workflow working, you often need to run it again with... [more](#)

Outcomes of the pro-iBiosphere workshop 4 on Business Models

10.12.2013

A pro-iBiosphere workshop on evaluation of business models took place on the 10th of October 2013 at the Botanischer Garten und Botanisches Museum (FUB-BGBM) in Berlin, Germany. It was attended by project partners and four external experts. The workshop was split into two sessions, each divided into smaller working groups.... [more](#)

pro-iBiosphere Meeting 4: Evaluation of the meeting

05.12.2013

From Oct 8-10th 2013 the 4th pro-iBiosphere meeting took place at the Botanic Garden and Botanical Museum Berlin-Dahlem. In total, 87 participants from 15 countries attended the 4 workshops held on: 1. 8 Oct Workshop 1 (M4.1): How to improve technical cooperation and interoperability at the e-infrastructure level (FUB-BGBM). For results... [more](#)

Improving the technical cooperation and interoperability at the e-infrastructure level

05.12.2013

A pro-iBiosphere workshop on "How to improve technical cooperation and interoperability at the e-infrastructure level" was held at the Botanic Garden and Botanical Museum Berlin-Dahlem (BGBM) on October 8 2013. A total of 22 participants were invited to attend the workshop, representing a wide range of international biodiversity-related institutions and... [more](#)

pro-iBiosphere project highlighted at the ICT2013 event

03.12.2013

The most visible forum for ICT research and innovation in Europe "ICT2013: Create, Connect and Grow", took place on the 6-8th of November 2013 in Vilnius (Lithuania). The event consisted of c.250 sessions and 200 exhibitors; and brought together lead thinkers and people driving European ICT research and innovation. A... [more](#)

Vibrant workshop

27.11.2013

Vibrant workshop for users of the EDIT Platform for Cybertaxonomy The Vibrant workshop for users of the EDIT Platform for Cybertaxonomy was held from 11-13 November at the Botanic Garden and Botanical Museum Berlin-Dahlem (FUB-BGBM). The aim of the workshop was to explain the new Taxonomic Editor, the Data Portals and... [more](#)

Novel ideas and tools from pro-iBiosphere presented at the TDWG 2013 Conference

27.11.2013

Lyubomir Penev, Gregor Hagedorn, Soraya Sierra The latest Biodiversity Information Standards (TDWG) conference took place in Florence, Italy on October 28 – November 1, 2013. The pro-iBiosphere project was well represented in the conference with 8 participants. Consortium members contributed especially to the discussions about: the future use of identifiers re-establishing the Structured Descriptive... [more](#)

The pro-iBiosphere project announces its final event

27.11.2013

Please mark your calendar for the pro-iBiosphere final event! The final meeting will take place on the 10-12th of June 2014 in Belgium. The National Botanic Garden of Belgium (NBGB) will be host to this meeting and has opened the doors of the Bouchout Castle for this event. During the meeting you will... [more](#)

TESS, a forerunner to pro-iBiosphere

11.11.2013

Prof. Robert E. Kenward [1] A project to design a Transactional Environmental Support System (www.tess-project.eu) preceded pro-iBiosphere during 2008-2011. TESS, of which partners were mostly linked to the Sustainable Use and Livelihoods Group of IUCN (www.iucn.org/SULi), recognised the need for information on species and habitats to inform policy, but also at... [more](#)

Toward best practices for sharing and maintaining biodiversity data - GBIF Finland (FinBIF)

06.11.2013

Hanna Koivula[1] The Finnish Museum of Natural History is a focal point for GBIF Finland (FinBIF). For the past few years we have been constructing a national research infrastructure to serve not only data needs of biodiversity researchers, but also environmental services. The aim is to fulfill many different data requirements... [more](#)

Flora of the Guianas: Toward an Online Platform

06.11.2013

Sylvia Mota de Oliveira[1] The three countries forming the Guianas - Guyana, Suriname and French Guiana - have more than 80% of their political territories covered by pristine Amazon forest. The Flora of the Guianas book series provides taxonomic treatments and associated data on plant species occurring in the region. At... [more](#)

Next-generation global e-infrastructure for taxon names registry

04.11.2013

Zookeys, the first zoological journal to introduce automatic registration in ZooBank The latest issue of ZooKeys - no. 346 - has been automatically registered in ZooBank on its day of publication last Friday. This marks the successful deployment of an automated registration-to-publication pipeline for taxonomic names for animals. The innovative workflow... [more](#)

The cyber-centipede: From Linnaeus to big data

29.10.2013

Taxonomic descriptions, introduced by Linnaeus in 1735, are designed to allow scientists to tell one species from another. Now there is a new futuristic method for describing new species that goes far beyond the tradition. The new approach combines several techniques, including next generation molecular methods, barcoding, and novel computing... [more](#)

Advanced Open Access publishing model

24.10.2013

The Biodiversity Data Journal goes beyond the basics of the Gold Open Access. There are two main modes of open access publishing – Green Open Access, where the author has the right to provide free access to the article outside the publisher's web site in a repository or on his/her own... [more](#)

New framework to deliver biodiversity knowledge

02.10.2013

Global Biodiversity Informatics Outlook sets out key steps to harness IT and open data to inform better decisions. Copenhagen, Denmark – A new initiative launched today (2 Oct) aims to coordinate global efforts and funding to deliver the best possible information about life on Earth, and our impacts upon it. The Global... [more](#)

Commission launches pilot to open up publicly funded research data

Iliyana Kuzmova 16.12.2013



Today, 16/12/2013, the European Commission announced the launch of a new Pilot on Open Research Data in Horizon 2020, to ensure that valuable information produced by researchers in many EU-funded projects will be shared freely. Researchers in projects participating in the pilot are asked to make the underlying data needed to validate the results presented in scientific publications and other scientific information available for use by other researchers, innovative industries and citizens. This will lead to better and more efficient science and improved transparency for citizens and society. It will also contribute to economic growth through open innovation. For 2014-2015, topic areas participating in the Open Research Data Pilot will receive funding of around €3 billion.

The Commission recognises that research data is as important as publications. It therefore announced in 2012 that it would experiment with open access to research data (see [IP/12/790](#)). The Pilot on Open Research Data in Horizon 2020 does for scientific information what the Open Data Strategy does for public sector information: it aims to improve and maximise access to and re-use of research data generated by projects for the benefit of society and the economy.

The Pilot involves key areas of Horizon 2020:

- Future and Emerging Technologies
- Research infrastructures – part e-Infrastructures
- Leadership in enabling and industrial technologies – Information and Communication Technologies
- Societal Challenge: Secure, Clean and Efficient Energy – part Smart cities and communities
- Societal Challenge: Climate Action, Environment, Resource Efficiency and Raw materials – with the exception of topics in the area of raw materials
- Societal Challenge: Europe in a changing world – inclusive, innovative and reflective Societies
- Science with and for Society

Neelie Kroes, Vice-President of the European Commission for the Digital Agenda said "We know that sharing and re-using research data holds huge potential for science, society and the economy. This Pilot is an opportunity to see how different disciplines share data in practice and to understand remaining obstacles."

Commissioner Máire Geoghegan-Quinn said: "This pilot is part of our commitment to openness in Horizon 2020. I look forward to seeing the first results, which will be used to help set the course for the future."

Projects may opt out of the pilot to allow for the protection of intellectual property or personal data; in view of security

concerns; or should the main objective of their research be compromised by making data openly accessible.

The Pilot will give the Commission a better understanding of what supporting infrastructure is needed and of the impact of limiting factors such as security, privacy or data protection or other reasons for projects opting out of sharing. It will also contribute insights in how best to create incentives for researchers to manage and share their research data.

The Pilot will be monitored throughout Horizon 2020 with a view to developing future Commission policy and EU research funding programmes.

The pro-iBiosphere project highlighted at the 6th EU-AU Cooperation Forum on ICT in Addis Ababa

Stephanie Morales 13.12.2013



The [6th Africa-EU Cooperation Forum on ICT](#) took place on December 2-3 2013 at the [African Union](#) Conference Center in Addis Ababa, Ethiopia under the aegis of the European Commission and the African Union Commission.

The event was organised by the EU [EuroAfrica-P8 project](#) on the occasion of the 50th anniversary of the African Union among the African ICT week.

During 20 sessions, 300 participants had the opportunity to share knowledge and explore the possibilities for cooperation in the framework of the Joint Africa-EU Strategic Partnership ([JAES](#)).

This meeting offered opportunities to the pro-iBiosphere project to promote its activities through the dissemination of project brochures and to network with potential stakeholders from Africa such as representatives from the Ethiopian agriculture portal on the occasion of the session on ICT for Agriculture on December 3.



Working, not drinking, at the Taverna

Iliyana Kuzmova 10.12.2013



Data analysis is a complicated and time-consuming process. Like a craftsman, you require a set of tools that source, reformat, merge and analyse data. Using these tools manually in a workflow can take weeks. Then, when you finally get the workflow working, you often need to run it again with a new set of inputs and parameters. What if there were a piece of software that could couple all these tools together and then run it all over again at a click of a button?. This is what [Taverna Workbench](#) does. Taverna, changes a time-consuming job with multiple tools into a single machine that does all the work seamlessly.

Taverna Workbench is one of the tools that supports the [BioVeL](#) project with their stated aim of creating a "virtual e-laboratory that supports research on biodiversity issues". Taverna, by itself, is like a conductor without an orchestra. The power of Taverna is in the flexible coupling together of web-services, scripts and all kinds of processing engines to create workflows. For example, an obvious use case is the coupling together of the webservice from GBIF with a niche model engine and rerunning of the workflow using different projections of future climate change. However, Taverna can be used to simplify the processing of practically any digital data. Many ecologists use R as their primary statistical software. R can be run from within Taverna, but Taverna helps you couple its running to pre- and post-processing so that it can be run more easily.

Taverna is widely used among phylogeneticists and bioinformaticians, but other disciplines are rapidly adopting it. Another, unique and powerful feature of Taverna is that people can share, distribute, and collaborate on their workflows. On the website [myExperiment.org](#) scientists post their workflows for others to use, criticise and improve upon. The website works like a social network, enabling users to create groups, "like" favorite workflows and exchange ideas. You could spend the time to program your own links between services, but Taverna lets you do this easily without sacrificing innovation and adaptability.

One of the important features of Taverna is the seamless way it allows users to use webservices. There is a growing list of webservices for biodiversity from organisations such as GBIF, EOL and EU-Brazil OpenBio. One of the big issues with webservices is that they are, almost by definition, invisible to human users. Therefore, how do you find out that they exist?. This is where [biodiversitycatalogue.org](#) comes in. It allows scientists to discover webservices, but also describes how they work and their input and output formats. The [pro-iBiosphere project](#) has helped to improve the catalogue and will help to set priorities for future development. It now [recommends the use of the Biodiversity Catalogue](#) as central service registration facility for a future Open Biodiversity Knowledge Management System.

Taverna is a relatively new system to the Biodiversity community and through the BioVeL project its user-base is growing rapidly. Furthermore, people are finding new uses for it all the time.

On 5th December 2013, BioVeL organised a workshop on Taverna at which pro-iBiosphere was represented. One potential use-case that is of interest to pro-iBiosphere is in the automated markup of text. Some aspects of automated markup are common to many texts, such as the identification of scientific names. On the other-hand there are other aspects that are specific to particular texts, such as the identification of treatment boundaries and language specific features. Taverna may be used to link generic services with custom scripts to significantly reduce the time it takes to markup text. Workflows could be created for one particular publication and then tweaked to work for another.

The possibilities of Taverna are almost limitless. It is just the glue, and you decide what you stick together. You might think the Taverna sounds like a quiet place for a drink, whereas, it is really the factory floor of data processing.

By Quentin Groom (National Botanic Garden of Belgium)

Outcomes of the pro-iBiosphere workshop 4 on Business Models

Stephanie Morales 10.12.2013



A pro-iBiosphere workshop on evaluation of business models took place on the 10th of October 2013 at the Botanischer Garten und Botanisches Museum ([FUB-BGBM](#)) in Berlin, Germany. It was attended by project partners and four external experts. The workshop was split into two sessions, each divided into smaller working groups. In the first session, a prioritised list of the partners' current products and services was drawn up, and the opportunities for, and threats to these were assessed. In the second session, the participants focussed on the services and activities that would comprise a future OBKMS (Open Biodiversity Knowledge Management System) and documented the constraints that might prevent the projected benefits of OBKMS from being realised.

The sessions have been very fruitful in terms of content (more than 20 matrix were made) and all participants (including external participants) have been very active during the whole day. Having external participants represented a real asset as they helped shaping the project vision more precisely while also demonstrating and confirming their interest in the OBKMS. Partners found these exercises very productive while taking time to step back and envision the future of the Consortium all together. This workshop is not the end of the exercise but only a milestone to agree on the various concepts, methodology and tools to be used to envision project sustainability and allow discussions among the partners. All in all, workshop objectives have been achieved.

The next steps of this workshop will be the release of an event report detailing the event outputs in presenting the project exploitation potential. We will keep you updated on the development of these activities.

For complementary information on the workshop (concept & objectives, agenda, participants list and presentations), visit the dedicated wiki page [here](#).



pro-iBiosphere Meeting 4: Evaluation of the meeting

Iliyana Kuzmova 05.12.2013

From Oct 8-10th 2013 the [4th pro-iBiosphere meeting](#) took place at the [Botanic Garden and Botanical Museum Berlin-Dahlem](#). In total, 87 participants from 15 countries attended the 4 workshops held on:

1. 8 Oct [Workshop 1](#) (M4.1): How to improve technical cooperation and interoperability at the e-infrastructure level (FUB-BGBM). For results from the workshop, see [here](#).
2. 8 Oct [Workshop 2](#) (M4.2): How to promote and foster the development & adoption of common mark-up standards & interoperability between schemas (PLAZI)
3. 9 Oct [Workshop 3](#) (M6.2.2): Workshop on user engagement and benefits (RBGK)
4. 10 Oct [Workshop 4](#) (M6.3.2): Towards sustainability towards service: Meeting to evaluate business models currently in use by partners and relevant non-partners (SIGMA)

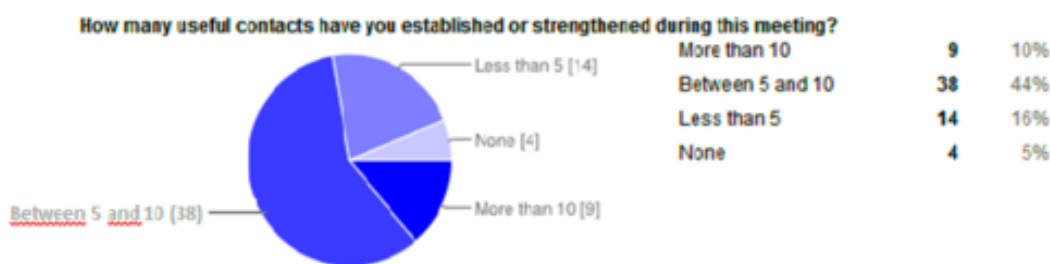
Networking

A questionnaire was sent to the participants of the meeting. A total of 87 persons answered the questionnaire. The meeting received an overall positive feedback. The agendas of the workshop and the possibility to network were the strongest attractions for attending. Half of the delegates were able to establish or strengthen 5-10 contacts during the event (see Figure 1). Delegates appreciated the discussions and welcoming atmosphere. Despite that, they mentioned that there could have been more time allocated for each workshop and each break for discussions.

Workshops

Participants expressed their preference to work in small groups with well-defined targets. The need for presentations was very low, provided that the workshops are well focused and give ample time for discussion.

30% of the participants judged the quality of the workshop as high, 31% as very good, 13% as acceptable and 1% as below expectations (see Figure 2). 50% of the participants are interested in attending other pro-iBiosphere events in the future and 64 out of 87 persons would recommend them to colleagues.



Caption: Contacts established by participants during the pro-iBiosphere meeting (first column; clockwise: 38% of the participants were able to make between 5 and 10 contacts, second column: 14% of the participants were able to establish less than 5

contacts, third column: 4% of the participants were not able to establish any contacts, fourth column: 9% of the participants were able to establish more than 10 contacts).

Improving the technical cooperation and interoperability at the e-infrastructure level

Iliyana Kuzmova 05.12.2013



A pro-iBiosphere workshop on "[How to improve technical cooperation and interoperability at the e-infrastructure level](#)" was held at the Botanic Garden and Botanical Museum Berlin-Dahlem ([BGBM](#)) on October 8 2013. A total of 22 participants were invited to attend the workshop, representing a wide range of international biodiversity-related institutions and e-infrastructures. The workshop focused on the establishment of two highly relevant interoperability aspects of: (i) a consistent space of stable identifiers for collection objects across European taxonomic institutions; and (ii) a central registry for biodiversity-related services.

In the workshop 8 different implementations of stable http-URI-based identifier systems in European- and US-based taxonomic institutions were positively evaluated. These implementations are an important outcome of the fruitful collaboration between pro-iBiosphere and the Information Science and Technology Commission (ISTC) of the Consortium of European Taxonomic Facilities (CETAF).

In addition, the workshop conducted a thorough analysis of the BiodiversityCatalogue (<https://www.biodiversitycatalogue.org/>) developed by the University of Manchester in the context of the EU 7th Framework project BioVeL. As a result, a detailed list of recommended improvements of the Catalogue was compiled and agreed on. The University of Manchester will use these recommendations for setting priorities when further developing the Catalogue. Detailed results from the workshop are available [here](#).

Authors: Anton Güntsch, Sabrina Eckert (FUB-BGBM)



pro-iBiosphere project highlighted at the ICT2013 event

Camille Torrenti 03.12.2013



The most visible forum for ICT research and innovation in Europe "[ICT2013](#): Create, Connect and Grow", took place on the 6-8th of November 2013 in Vilnius (Lithuania). The event consisted of c.250 sessions and 200 exhibitors; and brought together lead thinkers and people driving European ICT research and innovation. A total of 6.000 persons participated in the event, including researchers, innovators, entrepreneurs, industry representatives, and politicians.

ICT 2013 allowed participants to share best practices and experiences in big data management, and provided them an excellent opportunity to learn about the current state of ICT research in Europe and the new Horizon2020 Framework programme for Research and Innovation.

The pro-iBiosphere project was strongly represented during the event by means of an exhibition booth and a networking session co-organised with other EC-funded projects (i.e. ei4Africa, Chain-reds, e-Science Talk). The exhibition booth entitled 'e-Infrastructures at work and the future of research' showcased information from these four projects. Potential contacts were made with 20 stakeholders comprising projects on biodiversity data, EC-funded projects managing big data infrastructures (i.e. platforms, storage); and engineers specialised in semantic integration, enhancement, oncology, and autonomies.

The networking session on 'What does the future hold for e-Science and Big Data?' brought together researchers, data owners and service providers (including SMEs) to explore the future for e-science and how to deliver open access to data through Horizon2020. During this session, the pro-iBiosphere project (represented by Plazi) presented its vision and potential impact to the biodiversity community and beyond. The networking session led to better understanding of how e-infrastructures can solve scientific challenges. Additional information is available [here](#).



Vibrant workshop

Sabrina Eckert 27.11.2013

for users of the EDIT Platform for Cybertaxonomy



[Vibrant](#) workshop for users of the [EDIT Platform for Cybertaxonomy](#)

The Vibrant workshop for users of the EDIT Platform for Cybertaxonomy was held from **11-13 November** at the [Botanic Garden and Botanical Museum Berlin-Dahlem \(FUB-BGBM\)](#). The aim of the workshop was to explain the new Taxonomic Editor, the Data Portals and other components of the Platform to users, as well as to give an introduction to the software for structured taxonomic descriptions and keys, Xper2. The meeting brought together 40 participants representing a wide range of international biodiversity-related institutions and e-infrastructures ([Euro+Med](#) Plantbase, e-Floras, German Red Data Book editors, Pensoft publishers, Chinese Virtual Herbaria, Atlas Florae

Europaeae and more).

The workshop was split into two parts, 1.5 days for the EDITor and the data portals, 0.5 days for Xper2. For the EDITor and data portal workshop, 3 parallel working groups (1 in German and 2 in English language) were led by two developers and one or two taxonomists from the FUB-BGBM. One group was solely formed by Euro+Med Plantbase family editors and taxonomic experts. For Xper2 two parallel groups were led by four colleagues from Paris. To explain the new Taxonomic EDITor two virtual box images were created (the images had the software - data Portal, Taxonomic Editor and cdmserver preinstalled). One image included a simple dataset with 10 taxa and factual data from the Compositae tribe Cichorieae. The other image included a more complex dataset from Euro+MedPlantBase, to give the editors the opportunity to do the hands-on training with the data they are handling in the framework of this project.

All three groups were able to get extensive hands-on experience with the Taxonomic EDITor and to observe the direct interaction of the EDITor with the data portals. Feedback from the users was generally positive. Some of them expressed the need of a "light" version of the EDITor which makes it easier to use for less experienced users. Constructive criticism from the participants will help to improve functionalities of the EDITor, and a part of them will use the EDIT platform in the future for their projects. The introduction and hands-on training on Xper2 was also well received by the audience and yielded some fruitful feedback for the presenters of that part of the workshop.

<http://cybertaxonomy.eu/>



Novel ideas and tools from pro-iBiosphere presented at the TDWG 2013 Conference

Iliyana Kuzmova 27.11.2013

Lyubomir Penev, Gregor Hagedorn, Soraya Sierra

The latest Biodiversity Information Standards (TDWG) conference took place in Florence, Italy on October 28 – November 1, 2013. The pro-iBiosphere project was well represented in the conference with 8 participants. Consortium members contributed especially to the discussions about:

- the future use of identifiers
- re-establishing the Structured Descriptive Data standard in Semantic Web compatible RDF Format
- future developments in DarwinCore

Presentations given by consortium members during the meeting are available [here](#).

In order to make fundamental biodiversity data digital, open and re-usable, the pro-iBiosphere project has the vision of implementing an Open Knowledge Biodiversity System (OBKMS). The strong general current towards semantic technologies observed at the meeting, confirmed the need for this vision.

The conference is also to be remembered with the launch of the common automated registration model for higher plants,

fungi and animals (one of the four [pilots](#) being conducted by the pro-iBiosphere project). Currently registration of new taxa is being conducted "by hand" by the authors, journals or registries. Pensoft journals were first to develop and offer this new service, in a collaboration with the [International Plant name Index \(IPNI\)](#), [ZooBank](#) and [Index Fungorum](#). Detailed information on the automated registration pilot is available [here](#).

Preceding the TDWG conference, Pensoft and ZooBank organised a hackathon in Sofia (Bulgaria). The hackathon resulted in several, real-time tests of the pilot, based on the [TaxPub](#) XML schema.

After the hackathon, the teams of IPNI, Index Fungorum, ZooBank, Pensoft and Plazi came together at a brief workshop to discuss the potential of the [Taxonomic Concept Schema \(TCS\)](#) as a possible basic model that may be used in the future for registration of taxa in the three organismic domains.

The consortium strives towards having a more open pilot registration system available by the end of the pro-iBiosphere project. At present, IPNI is having discussions with other key partners such as the IAPT committee on registration on how to open up the automated registration model.

The pro-iBiosphere project announces its final event

Stephanie Morales 27.11.2013



Please mark your calendar for the pro-iBiosphere final event!

The [final meeting](#) will take place on the 10-12th of June 2014 in Belgium.

The National Botanic Garden of Belgium (NBGB) will be host to this meeting and has opened the doors of the Bouchout Castle for this event.

During the meeting you will have the opportunity to learn about outcomes of the project, be informed about relevant topics presented by invited speakers, follow training on state-of-the art tools, and do networking with colleagues.

We will keep you updated as our preparations develop.

In case of further questions please do not hesitate to contact us at pro-ibiosphere@sigma-orionis.com.

We look forward to seeing you in June!



TESS, a forerunner to pro-iBiosphere

Iliyana Kuzmova 11.11.2013



Prof. Robert E. Kenward [\[1\]](#)

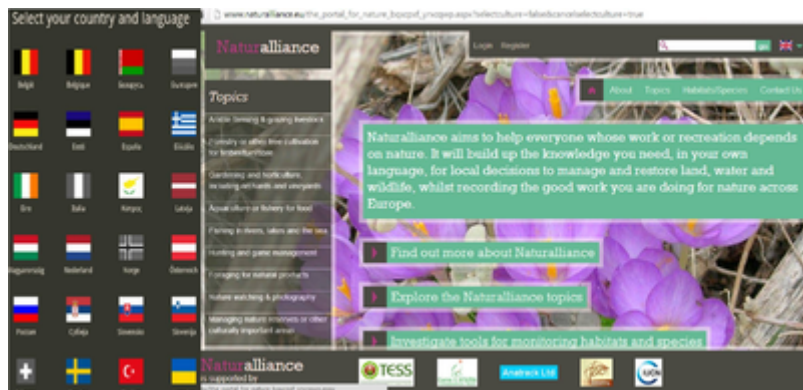
A project to design a Transactional Environmental Support System (www.tess-project.eu) preceded pro-iBiosphere during 2008-2011. TESS, of which partners were mostly linked to the Sustainable Use and Livelihoods Group of IUCN (www.iucn.org/SULi), recognised the need for information on species and habitats to inform policy, but also at local level for management of land and sustainable use. TESS surveyed information needs at EU level, national governments, local communities at lowest government administrative level and stakeholders. The surveys not only showed biodiversity information requirements in detail, but also found a "decision density" five magnitudes higher for local managers of land and species (in farms, forests, fisheries, hunting areas and nature reserves) than for statutory environmental assessments (SEA+EIA).

To make decisions that are good for nature, there is scope for internet exchange (transaction) of science-based decision support (from high level) for the detailed data needed to make those local decisions (but also needed for policy-making at higher level). For effective exchange, data needs to be open access, standardised in format and highly detailed (accurate and at high resolution), i.e. aspirations of pro-iBiosphere. Case studies showed that local communities were highly capable of collecting appropriate data, but that thousands of models for science-based decision support included only 10 suitable for local use, at most in two languages!

This information was of special relevance for one of the pro-iBiosphere workshops in October 2013, [Workshop on User Engagement and Benefits](#), showing in particular the importance of delivering Faunas, Floras and Mycotas for local people to record biodiversity. Identifying and reporting species needs to be made simple and multilingual for users other than experts, which could involve smart-phone apps for image-recognition and database-linking.

As part of its surveys, TESS established a portal (www.naturalliance.eu) to signpost relevant information and encourage collection by local interests. Naturalliance is advised by major international NGOs at European and Global level, is now in 23 languages and, to encourage sustainable use of land and biodiversity, and is working with national governments on a new system to encourage systematic local participation in data exchange. Experience with funding these systems, which could signpost a one-stop-shop for identification/reporting apps, or even host the downloading, was shared during the [meeting](#) on

[1] Vice-chair (Europe) of IUCN Sustainable Use and Livelihoods Specialist Group



Toward best practices for sharing and maintaining biodiversity data - GBIF Finland (FinBIF)

Iliyana Kuzmova 06.11.2013



Hanna Koivula[1]

The [Finnish Museum of Natural History](#) is a focal point for GBIF Finland (FinBIF). For the past few years we have been constructing a national research infrastructure to serve not only data needs of biodiversity researchers, but also environmental services. The aim is to fulfill many different data requirements for various kinds of surveys and assessments concerning biodiversity. Most of our data are openly available to anyone, but technically many data users have not been able to use our web services for retrieving the data they need. The main obstacle is that the services are very new and the awareness about exchange protocols and standards of the data is poor. This is why most of the data shared by FinBIF are still extracted manually upon request. Hence, one (still unreached) goal is to name and establish best practices for sharing and maintaining biodiversity data. This is done by applying selected parts, especially for logical architecture and communication, from Enterprise architecture. The method enables us to capture essential information of our research infrastructure under construction.

Building a research infrastructure in an extremely heterogeneous "funding environment" has set some serious impediments along the way. On the other hand, it has allowed us to play with quite adventurous architectural structures and innovative pilot cases. One of these new approaches is the use of semantic web features for making data storage more flexible and data themselves better discoverable.

Participating in the pro-iBiosphere workshop on "[How to improve technical cooperation and interoperability at the e-infrastructure level!](#)" that took place on the 8th of October 2013, in Berlin, gave me a good insight into existing European open biodiversity data and related services. Hands-on experience of listing our web services to the Biodiversity Catalogue triggered useful conversations on how to improve the interoperability and discoverability of the services, and gave participants take-home ideas. In the break-out discussion group we assessed the present set of metadata elements used in the Biodiversity Catalogue for describing individual services. Based on that evaluation, we made recommendations for

additional elements. I also obtained many good ideas for further developing the FinBIF services.

[1] Finnish Museum of Natural History - IT Specialist GBIF Node Manager for Finland, Email: hanna.koivula@helsinki.fi

Flora of the Guianas: Toward an Online Platform

Iliyana Kuzmova 06.11.2013



Sylvia Mota de Oliveira[1]

The three countries forming the Guianas - Guyana, Suriname and French Guiana - have more than 80% of their political territories covered by pristine Amazon forest. The Flora of the Guianas book series provides taxonomic treatments and associated data on plant species occurring in the region.

At present, the published fascicles cover around 25% of the total of species recorded in the Guianas, with a total of 2294 species treated only in series A: Phanerogams and around 1000 species in other series, covering Ferns and Fern allies, Bryophytes and Fungi. Beyond taxonomy, the Flora is a rich and well curated source of primary biodiversity data. To efficiently use this source and build it further, we want to move towards an online platform, where:

- Data extracted from other literature sources (Flora of Suriname, for instance) can be incorporated in a database and used as a basis for upcoming taxa treatments, increasing the rate of species descriptions;
- Data gathered in the Flora fascicles can be disclosed and made available in a structured and searchable manner, openly accessible to end-users;
- Interaction between Flora contributors, specialists and citizen scientists is promoted;
- Links between elements of the Flora and external content, such as collection and molecular databases, can be easily enabled;
- Updates of the taxonomic data can guarantee the quality and usability of the Flora in several other research fields.

To achieve these goals, the transference of the data from the printed copies of the Flora of the Guianas to such an online platform requires mark-up of text. The new workflow is being tested in a [Pilot](#) study within the pro-iBiosphere project, which will hopefully support the planning of a longer-term strategy for the complete Flora.

The first results of the pilot were presented and discussed during the pro-iBiosphere workshop on "[How to promote and foster the development & adoption of common mark-up standards & interoperability between schemas](#)" that took place on the 8th of October 2013, in Berlin. The workshop was very instructive and useful for the future strategy of the Flora of the Guianas. In this workshop, participants shared their experiences, and discussed together the most important issues/challenges that they have encountered while working on text mark-up. The heterogeneity of the group added to the discussion because participants were confronted with different levels of: 1) experience with mark-up tools; 2) data granularity; 3) consistency in the format of the content to be marked-up. Relevant topics discussed included the scaling of the mark-up activities using different tools, the costs and possible incentives for such an activity, the desired level of granularity of the mark-up, and best practices to be adopted by taxonomists and bioinformaticians.

[1] Editor of Flora of the. Guianas, Naturalis Biodiversity Center, the Netherlands. Contact:

Next-generation global e-infrastructure for taxon names registry

Iliyana Kuzmova 04.11.2013

Zookeys, the first zoological journal to introduce automatic registration in ZooBank

The latest issue of [ZooKeys](#) - no. 346 - has been automatically registered in [ZooBank](#) on its day of publication last Friday. This marks the successful deployment of an automated registration-to-publication pipeline for taxonomic names for animals. The innovative workflow was jointly funded by the EU FP7 funded project [pro-iBiosphere](#) and a U.S. National Science Foundation project to develop the Global Names Architecture (DBI-1062441).

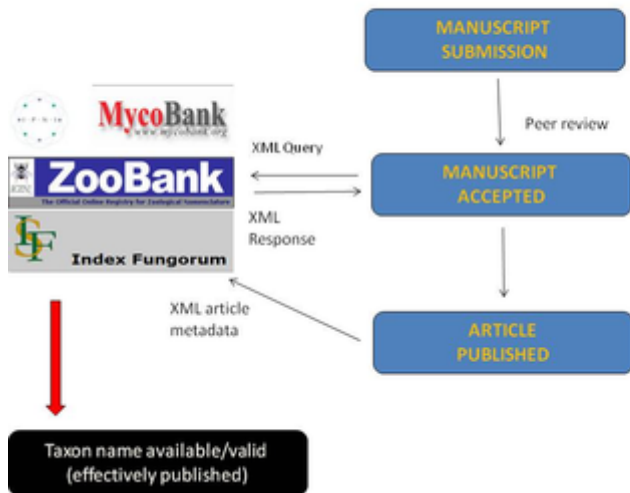
The process of post-publication recording and indexing of species names has a long tradition, in some cases dating as far back as the middle of 19th century. But now in the 21st century with the advance of modern technologies and the opportunity to publish taxonomic novelties online, the process of post-publication recording brought into focus the concept of automated pre-publication registration.

Why is this important? The proportion of 'turbo-taxonomic' papers describing hundreds of new species increases. Registration of hundreds of new species is an issue, however it is even more important that the final publication data of the pre-registered names are reported back to ZooBank on the day of publication.

Launched as an open access peer reviewed journal in 2008, to coincide and adopt from inception the International Code of Zoological Nomenclature changes for electronic publications, *ZooKeys* was the first journal to provide a mandatory in-house registration in ZooBank. Since 2008, it has contributed about one third of all names currently registered in ZooBank. With the adoption of the automated ZooBank registration, *ZooKeys* continues its mission to set novel trends in biodiversity publishing.

Implementation of automated workflows and invention of XML-based tools will facilitate the process of publication and dissemination of biodiversity information. It will pave the way for unification and streamlining the registration process, even more to building the next-generation e-infrastructure for a common global taxon names registry. Within the pro-iBiosphere project and in cooperation with [Plazi](#) that have created the [TaxPub](#) XML schema, an automated registration workflow for plants has already been established between the [International Plant Names Index](#) (IPNI) and the [PhytoKeys](#) journal, to be applied soon also for fungi between [Index Fungorum](#) and the journal [MycoKeys](#).

Automated registration



The cyber-centipede: From Linnaeus to big data

Iliyana Kuzmova 29.10.2013

Taxonomic descriptions, introduced by Linnaeus in 1735, are designed to allow scientists to tell one species from another. Now there is a new futuristic method for describing new species that goes far beyond the tradition. The new approach combines several techniques, including next generation molecular methods, barcoding, and novel computing and imaging technologies, that will test the model for big data collection, storage and management in biology. The study has just been published in the [Biodiversity Data Journal](#).

While 13,494 new animal species were discovered by taxonomists in 2012, animal diversity on the planet continues to decline with unprecedented speed. Concerned with the rapid disappearance rates scientists have been forced towards a so called 'turbo taxonomy' approach, where rapid species description is needed to manage conservation.

While acknowledging the necessity of fast descriptions, the authors of the new study present the other 'extreme' for taxonomic description: "a new species of the future". An international team of scientists from Bulgaria, Croatia, China, UK, Denmark, France, Italy, Greece and Germany illustrated a holistic approach to the description of the new cave dwelling centipede species *Eupolybothrus cavernicolus*, recently discovered in a remote karst region of Croatia. The project was a collaboration between [GigaScience](#), [China National GeneBank](#), [BGI-Shenzhen](#) and [Pensoft Publishers](#).

Eupolybothrus cavernicolus has become the first eukaryotic species for which, in addition to the traditional morphological description, scientists have provided a transcriptomic profile, DNA barcoding data, detailed anatomical X-ray microtomography (micro-CT), and a movie of the living specimen to document important traits of its behaviour. By employing micro-CT scanning in a new species, for the first time a high-resolution morphological and anatomical dataset is created - the '[cybertype](#)' giving everyone virtual access to the specimen.

This, most data-rich species description, represents also the first biodiversity project that joins the ISA (Investigation-Study-Assay) Commons, that is an approach created by the genomic and molecular biology communities to store and describe different data types collected in the course of a multidisciplinary study.

"Communicating the results of next generation sequencing effectively requires the next generation of data publishing" says Prof. Lyubomir Penev, Managing director of Pensoft Publishers. "It is not sufficient just to collect 'big' data. The real challenge comes at the point when data should be managed, stored, handled, peer-reviewed, published and distributed in a way that allows for re-use in the coming big data world", concluded Prof. Penev.

"Next generation sequencing is moving beyond piecing together a species genetic blueprint to areas such as biodiversity research, with mass collections of species in "metabarcoding" surveys bringing genomics, monitoring of ecosystems and species-discovery closer together. This example attempts to integrate data from these different sources, and through curation in BGI and *GigaScience's* GigaDB database to make it interoperable and much more usable," says Dr Scott Edmunds from BGI and Executive Editor of *GigaScience*.

Additional information:

Pensoft and the Natural History Museum London have received financial support by the EU FP7 projects [ViBRANT](#) and [pro-iBiosphere](#). The China National GeneBank (CNGB) and *GigaScience* teams have received support from the BGI. The DNA barcodes were obtained through the International Barcode of Life Project supported by grants from NSERC and from the government of Canada through Genome Canada and the Ontario Genomics Institute.

Original Sources:

Stoev P, Komeri?ki A, Akkari N, Shanlin Liu, Xin Zhou, Weigand AM, Hostens J, Hunter CI, Edmunds SC, Porco D, Zapparoli M, Georgiev T, Mietchen D, Roberts D, Faulwetter S, Smith V, Penev L (2013) *Eupolybothrus cavernicolus* Komeri?ki & Stoev sp. n. (Chilopoda: Lithobiomorpha: Lithobiidae): the first eukaryotic species description combining transcriptomic, DNA barcoding and micro-CT imaging data. *Biodiversity Data Journal* 1: e1013. DOI: [10.3897/BDJ.1.e1013](https://doi.org/10.3897/BDJ.1.e1013)

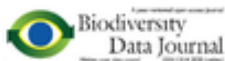
Edmunds SC, Hunter CI, Smith V, Stoev P, Penev L (2013) Biodiversity research in the "big data" era: *GigaScience* and Pensoft work together to publish the most data-rich species description. *GigaScience* 2:14 doi:[10.1186/2047-217X-2-14](https://doi.org/10.1186/2047-217X-2-14)

Watch the 3D cybertype video: http://www.youtube.com/watch?v=vqPuwKG8hE4&feature=em-upload_owner



Advanced Open Access publishing model

Iliyana Kuzmova 24.10.2013



The Biodiversity Data Journal goes beyond the basics of the Gold Open Access

There are two main modes of open access publishing – Green Open Access, where the author has the right to provide free access to the article outside the publisher's web site in a repository or on his/her own website, and Gold Open Access, where articles are available for free download directly from the publisher on the day of publication.

Opening of content and data, however does not necessarily mean "easy to discover and re-use". The [Biodiversity Data](#)

[Journal](#) proposed the term "Advanced Open Access" to describe an integrated, narrative (text) and data publishing model where the main goal is to make content "re-usable" and "interoperable" for both humans and computers.

To publish effectively in open access, it is not sufficient simply to provide PDF or HTML files online. It is crucial to put these under a reuse-friendly license and to implement technologies that allow machine-readable content and data to be harvested and collated into a big data pool.

The Advanced Open Access means:

- Free to read
- Free to re-use, revise, remix, redistribute
- Easy to discover and harvest
- Content automatically summarised by aggregators
- Data and narrative integrated to the widest extent possible
- Human- and computer-readable formats
- Community-based, pre- and post-publication peer-review
- Community ownership of data
- Free to publish or at low cost affordable by all

BDJ shortens the distance between "narrative" (text) and "data" publishing. Many data types, such as species occurrences, checklists, measurements and others, are converted into text from spreadsheets for better readability by humans. Conversely, text from an article can be downloaded as structured data or harvested by computers for further analysis.

"Open access is definitely one of the greatest steps in scientific communication comparable to the invention of the printing technology or the peer-review system. Great but not sufficient!" said Prof. Lyubomir Penev, founder of [Pensoft Publishers](#) and the *Biodiversity Data Journal*. "We need to switch the focus already from making content 'available for free download' to being discoverable and extractable. Such re-usability multiplies society's investment in science".

###

Additional information:

The *Biodiversity Data Journal* is designed by Pensoft Publishers and was funded in part by the European Union's Seventh Framework Program (FP7) project [ViBRANT](#).

Source: Smith V, Georgiev T, Stoev P, Biserkov J, Miller J, Livermore L, Baker E, Mietchen D, Couvreur T, Mueller G, Dikow T, Helgen K, Frank J, Agosti D, Roberts D, Penev L (2013) Beyond dead trees: integrating the scientific process in the *Biodiversity Data Journal*. *Biodiversity Data Journal* 1: e995. DOI: [10.3897/BDJ.1.e995](https://doi.org/10.3897/BDJ.1.e995)

New framework to deliver biodiversity knowledge

Iliyana Kuzmova 02.10.2013



Global Biodiversity Informatics Outlook sets out key steps to harness IT and open data to inform better decisions

Copenhagen, Denmark – A new initiative launched today (2 Oct) aims to coordinate global efforts and funding to deliver the best possible information about life on Earth, and our impacts upon it.

The Global Biodiversity Informatics Outlook sets out a framework to harness the immense power of information technology and an open data culture to gather unprecedented evidence about biodiversity and to inform better decisions.

The framework is outlined in a [document](#) and [website](#) entitled *Delivering Biodiversity Knowledge in the Information Age*,

inviting policy makers, funders, researchers, informatics specialists, data holders and others to unite around four key focus areas where progress is needed.

The focus areas, each consisting of several specific components, are:

- **Culture** – promoting practices and infrastructure for sharing data, using common standards and persistent archives, backed up by strong policy incentives and a community of willing specialists;
- **Data** – addressing the need to transform all data about species, past and present, into usable and accessible digital formats; from historic collections and literature to citizen science observations, remote sensors and gene sequencing;
- **Evidence** – organizing and assessing data from all sources to provide clear, consistent views giving them context; including taxonomic organization, integrated occurrences in time and space, capturing information about species interactions, and improving data quality through collaborative curation; and
- **Understanding** – building models from recorded measurements and observations to support data-driven research and evidence-based planning, including predictive tools, better visualization and feedbacks to prioritize new data capture.

The document is being promoted through a number of upcoming events this month, including the [Governing Board of the Global Biodiversity Information Facility](#) and the [Subsidiary Body on Scientific, Technical and Technological Advice of the Convention on Biological Diversity](#) (CBD SBSTTA) where it forms part of the discussion on meeting global targets to end biodiversity loss.

The framework arose from the [Global Biodiversity Informatics Conference](#) which gathered around 100 experts in Copenhagen in July, 2012, to identify critical questions relating to biodiversity and tools needed answer them. Workshop leaders at that conference went on to draw up and author the current document.

The Global Biodiversity Informatics Outlook includes examples of projects and initiatives contributing to its objectives, and the accompanying website www.biodiversityinformatics.org invites feedback from others wishing to align their own activities to the framework.

A deck of slides for presentations about GBIO is available at <http://www.slideshare.net/GBIF/global-biodiversity-informatics-outlook>