

Milestone MS261, MS262 and MS272

Version: 1.0
Date: 2013-06-04
Author: MRAC_UEF
Document reference: Milestone_MS261 MS262 MS272



Initial Informatics Workshop:
MS261: Helpdesk online and technical coordination unit established (M05)
MS262: Introduction to relevant projects and technologies to training programmes and helpdesk (M7)
and
MS272: First meeting of Informatics and Data Standards Task Force (M8)

STATUS: FINAL

Project acronym: EU BON
Project name: EU BON: Building the European Biodiversity Observation Network
Call: ENV.2012.6.2-2
Grant agreement: 308454
Project Duration: 01/12/2012 – 31.05.2017 (54 months)
Co-ordinator: MfN, Museum für Naturkunde - Leibniz Institute for Research on Evolution and Biodiversity, Germany

Partners: UTARTU, University of Tartu, Natural History Museum, Estonia
UEF, University of Eastern Finland, Digitisation Centre, Finland
GBIF, Global Biodiversity Information Facility, Denmark
UniLeeds, University of Leeds, School of Biology, UK
UFZ, Helmholtz Centre for Environmental Research, Germany
CSIC, The Spanish National Research Council, Doñana Biological Station, Spain
UCAM, University of Cambridge, Centre for Science and Policy, UK
CNRS-IMBE, Mediterranean Institute of marine and terrestrial Biodiversity and Ecology, France
Pensoft, Pensoft Publishers Ltd, Bulgaria
SGN, Senckenberg Gesellschaft für Naturforschung, Germany
VIZZUALITY, Vizzuality S.L., Spain
FIN, FishBase Information and Research Group, Inc., Philippines
HCMR, Hellenic Centre for Marine Research, Greece
NHM, The Natural History Museum, London
BGBM, Botanic Garden and Botanical Museum Berlin-Dahlem, Germany
UCPH, University of Copenhagen: Natural History Museum of Denmark, Denmark
RMCA, Royal Museum of Central Africa, Belgium
PLAZI, Plazi GmbH, Switzerland
GlueCAD, GlueCAD Ltd. – Engineering IT, Israel
IEEP, Institute for European Environmental Policy, UK
INPA, National Institute of Amazonian Research, Brazil
NRM, Swedish Museum of Natural History, Sweden
IBSAS, Slovak Academy of Sciences, Institute of Botany, Slovakia
EBCC-CTFC, Forest Technology Centre of Catalonia, Spain
NBIC, Norwegian Biodiversity Information Centre, Norway
FEM, Fondazione Edmund Mach, Italy
TerraData, TerraData environmetrics, Monterotondo Marittimo, Italy
EURAC, European Academy of Bozen/Bolzano, Italy
WCMC, UNEP World Conservation Monitoring Centre, UK

This project is supported by funding from the specific programme 'Cooperation', theme 'Environment (including Climate Change)' under the 7th Research Framework Programme of the European Union.



EU BON

EU BON: Building the European Biodiversity Observation Network
Project no. 308454

Large scale collaborative project

MS261

Helpdesk online and technical coordination unit established




MS262

Introduction to relevant projects and technologies to training programmes and helpdesk

MS272

First meeting of Informatics and Data Standards Task Force

Milestone number	MS261, MS262 and MS272
Milestone name	-Helpdesk online and technical coordination unit established -Introduction to relevant projects and technologies to training programmes and helpdesk and -First meeting of Informatics and Data Standards Task Force
WP no.	WP2
Lead Beneficiary (full name and Acronym)	University of Eastern Finland, Digitisation Centre, UEF Royal Museum of Central Africa, MRAC
Nature	Written Report
Delivery date from Annex I (proj. month)	2013-05-31 and 2013-06-30
Delivered	yes
Actual forecast delivery date	2013-05-31
Comments	

Name of the Authors	Name of the Partner	Logo of the Partner
Hannu Saarenmaa	University of Eastern Finland	 UNIVERSITY OF EASTERN FINLAND
Patricia Mergen	MRAC	 ROYAL MUSEUM FOR CENTRAL AFRICA TERVUREN
Kim Jacobsen	MRAC	 ROYAL MUSEUM FOR CENTRAL AFRICA TERVUREN

In case the report consists of the delivery of materials (guidelines, manuscripts, etc)

Delivery name	Delivery name	From Partner	To Partner

Introduction

The current report concerns three milestones:

- MS 261: Helpdesk online and technical coordination unit established (due M5)
- MS 262: Introduction to relevant projects and technologies to training programmes and helpdesk services (due M7)
- MS 272: First meeting of Informatics and Data Standards Task Force (due M8)

The development of MS 261 and MS 262 was largely initiated and influenced by discussions held during MS 272 (First meeting of Informatics and Data Standards Task Force), hence the decision to merge these three milestones in the current report.

Progress towards objectives

The first EU BON Informatics Workshop (**MS 272** *First meeting of Informatics and Data Standards Task Force*) took place in Trondheim, Norway from 29 to 31 May 2013 and was hosted by EU BON partner *Norwegian Biodiversity Information Centre*. The main focus of the event was the discussions of data standards and informatics architecture to be followed by the EU BON project. The workshop also focussed on highlighting linkages to GEOSS Common Infrastructure and the GEO BON, plans and operations of the EU BON Helpdesk and the first assembly of the global Informatics Task Force (ITF) of EU BON. The agenda and participants are detailed in Annex I; the minutes of the meeting are provided in Annex II.

Presentations are available from the EU BON Helpdesk website (<http://eubon.cybertaxonomy.africamuseum.be/file-browser>). Presentations are available from the internal ICP library under Workpackage 2 (see meetings): <http://www.eubon.eu/library/>

Members of the international Informatics Task Force lead the discussions on the first day when presentations were given to set the scene for the meeting. The second day focussed on work planning with detailed discussions of the various EU-BON tasks. The respective Task Leaders introduced MS211 (CSIC) and MS221 (GBIF), and writing tasks were assigned to the various partners. Helpdesk operations, tools for data sharing and training programmes were described by the Task Leader for Tasks 2.3, 2.6 and 2.8 (MRAC: Patricia Mergen). In the afternoon, meeting participants split into three focus groups (track 1, track 2 and track 3; see Annex I) for 2 hours, before reporting back and allowing time for a joint discussion of the main conclusions.

On the third day, agreements on the direction of the informatics developments were made, and action lists were formulated. Detailed minutes of the discussions are provided in Annex II.

Agreements and action list

A draft outline of the architectural design (MS211) and a draft outline of the data standards (MS221), were both discussed during the first meeting of the Data Standards Task Force (MS272). The Task Leader will prepare the first draft of MS211 for review by end of September 2013. The Partners will then comment by end of October 2013. The second draft will be circulated by the Task Leader by the middle of December 2013.

For MS221, it was agreed that the Task Leader will distribute the outline of the review paper, where writing assignments have been marked. The writing will be done by partners by

middle of September 2013. The Task leader will then distribute a combined first draft by end of September 2013, and further writing will be agreed on.

Regarding **MS272** (ITF meetings), it was agreed that the ITF will meet from now on tele-conference on a quarterly basis. The next meeting will be in the beginning of September 2013. A following meeting is scheduled during the TDWG meeting in October 2013, with follow-up correspondence in the context of the D2.1.

It was agreed to contribute to the goal of GEO BON to become operational by 2015. It was understood that this will require speeding up the timing of some activities from what was determined in the DoW. Two products will be developed, in particular (1) a basic infrastructure and (2) a novel use case to support GEO BON and EU policies.

For the infrastructure product (1), the following statement in the DoW was used as a template for discussions: “*EU BON will provide integration between social networks of science and policy and technological networks of interoperating IT infrastructures, resulting in a **new open-access platform for sharing biodiversity data and tools, and greatly advance biodiversity knowledge in Europe.***”

It was agreed that a new platform integrating primary biodiversity data and ecosystem data is needed, and should be built upon existing solutions. This platform is described in further detail in D2.1, the EUBON Portal is a part of this platform. The new platform would technically consist of the solutions of the DataONE network, augmented with semantic mediation for primary biodiversity data. EU BON Partners who wish to collaborate should implement DataONE Member Nodes, which will be populated with data from the networks and sites where each of the Partners are involved. A DataONE coordinating node in Europe may be established towards 2015, and data will later be integrated via the EU BON Portal. (Relevance to DoW: D2.2, D2.4 and MS231)

MS231: Specifications of data sharing tools (M14); D2.2: Data sharing tools (M39); D2.4: Report and assessment of training activities and final versions of training manuals (M45);

For the second product (2), the novel use case, it was agreed to join and support the GEO BON Working Group pilot project on automating the data flows for the Essential Biodiversity Variables (EBVs). This will require speeding up the EU BON registry and Portal implementation and linking with GEOSS Common Infrastructure. The pilot project could be implemented as part of the GEOSS 7th Architecture Implementation Pilot (AIP-7) for which a specification is needed in February 2014, the working pilot by late 2014, and a demonstration video soon after. This is relevant as it contributes towards IPBES and EU initiatives such as SEIS. (Relevance to DoW: D2.3, D2.4, MS241, MS251)

MS241: Specifications for registry and metadata catalogue (M27); MS251: Specification for the EU BON portal (M27); D2.3: Registry and metadata catalogue (M 39); D2.4: Report and assessment of training activities and final versions of training manuals (M45);

Following the Trondheim meeting, further progress has been made regarding the EU-BON helpdesk (**MS 261 Helpdesk online and technical coordination unit established**). The helpdesk aims to:

1. Support developments made in other WPs
2. Enable the review of design documents, by the testing developments and providing feedback to the developers

3. Promoting activities of national biodiversity observation networks (BONs) and other users
4. Assisting BONs to install and use tools that foster interoperability mechanisms (Task 2.3)
5. Promoting open access
6. Assisting the registration of EU-Bon services at the GEOSS Common Infrastructure (GCI)
7. Provide an interface to enable posting the same content to multiple domains, thus helping to populate the portal with content from other WPs and partners
8. Provide access to the Biodiversity Tools Repository BDTracker.net
9. Provide access to a repository of relevant documents
10. Enable easy access to expert knowledge

This centralization of information will minimize duplication of efforts and the creation of multiple access points. The helpdesk is currently not operating at full potential (<http://eubon.cybertaxonomy.africamuseum.be/>). A sequence of actions has been recommended and will be implemented over the next few months by MRAC in consultation and close collaboration with the partners for this task UEF and Pensoft. Estimated date of completion is 31/03/2014. These actions form the basis for the implementation of **MS262** (*Introduction to relevant projects and technologies to training programmes and helpdesk*), including:

- Enabling the registration of observation data in the GEOSS registry: a step-by-step guide will be created and made available on the helpdesk (reference: "GEO BON Detailed Implementation Plan" and "Principles of information architecture").
- DataONE training activities will be cited on the helpdesk, with updated links and RSS feed. A link will be made with the DataONE website on how to become a DataONE Member node. A DataONE Member node will be set-up for demo and training purposes.
- The promotion of open access will be fostered through visible linkage to the GEOSS Data Sharing Principles and highlighting that data sharing means registering it with a data registry service. A list of links to biodiversity portals and data registries, i.e. GBIF, KNB/DataONE, GEOSS; BiodiversityCatalogue of BioVeL, EUMON, LTER-Europe DEIMS will also be provided
- Similarly a list of links to available data sharing tools and information on the modalities of data sharing will be provided.
- Best practice documents created by the above-mentioned networks will also be uploaded to the helpdesk.

Annex I

EU BON Initial Informatics Workshop Trondheim, Norway, 29-31 May 2013

AGENDA and ATTENDEES

The first EU BON Informatics Workshop will take place in Trondheim, Norway from 29 to 31 May 2013, and will be hosted by EU BON partner *Norwegian Biodiversity Information Centre*. The event aims at discussing the data standards and informatics architecture that will be followed by the EU BON project. The workshop will also focus on highlighting linkages to GEOSS Common Infrastructure and the GEO BON. Among the agenda on the list are also a presentation of the plans and operations of the EU BON Helpdesk as well as the first assembly of the global Informatics Task Force (ITF) of EU BON.



Figure 1: Participants of the EU BON Initial Informatics Workshop (Trondheim, 29-31 May 2013)

Venue

NINA (and NBIC) Office and Conference Centre at Høgskoleringen 9, Trondheim, Norway.
<http://kart.gulesider.no/m/ns3g0>

Agenda

29 May - Wednesday afternoon: Setting the scene

1100 Registration

- 1200 Lunch in the restaurant of the Conference Centre, Høyskoleringen 9
 1300 Welcome by local organizers, the Norwegian Biodiversity Information Centre
 1310 Welcome by EU BON WP2 leader (Hannu Saarenmaa)
 1320 Round of introductions

Getting to know GEOSS and GEO BON

- 1330 GEOSS Common Infrastructure (GCI) and linking community resources with it (Stefano Nativi)
 1410 GEO BON history, aims for 2015, and current priorities such as Essential Biodiversity Variables, and activities of the GEO BON Working Group 8 “Data Integration and Interoperability” (Wim Hugo)
 1450 Coffee

What can EU BON learn from and build on work done elsewhere

- 1520 Overview of activities of the DataONE project in the USA (Mark Schildhauer)
 1600 EnvEurope project and LTER Europe: Metadata and accessibility to data resulting from long term ecological research and monitoring (Johannes Peterseil)
 1620 LifeWatch and Creative-B (Wouter Los)
 1640 BioVeL - Biodiversity Virtual e-Laboratory (Jonathan Giddy)
 1700 ViBRANT: Small-data holders as contributors into bigger data pools through standardised and streamlined environments (Dimitrios Koureas)
 1720 INSPIRE data specification on Annex III, 19 Species distribution (Nils Valland)
 1740 A few thoughts on the role of EU BON in supporting the EU in its Biodiversity Strategy to 2020 (Gregoire Dubois, talk by teleconference link)
 1800 Summary and discussion
 1830 Return to Rica Bakklundet Hotel
 2000 Dinner hosted by NTNU/NBIC

30 May - Thursday: Planning of work

- 0900 Overview of the aims of the EU BON project and the description of the work of the EU BON Work Package 2 “Data Integration and Interoperability” (Hannu Saarenmaa)
 0915 Task 2.7. - Global coordination through the EU BON Informatics Task Force (ITF): Communications needed to make this happen (Discussion)

Detailed discussion and group work on the EU BON tasks; overviews by Task Leaders

- 0930 Task 2.1 - Design of information architecture for EU BON (Carlos Rodriguez)
 0940 Task 2.2 - Improving data standards and interoperability (Éamonn Ó Tuama)
 0950 Task 2.3 - Tools for data sharing (Patricia Mergen)
 1000 Coffee
 1030 Task 2.4 - Metadata registry and catalogue (Éamonn Ó Tuama)
 1040 Task 1.1 - Assessment and evaluation of biodiversity data sources (Urmas Kõljalg)
 1050 Task 2.5 - European Biodiversity Portal (Carlos Rodriguez)
 1100 Task 2.6 - Technical support and helpdesk (Patricia Mergen)
 1110 Task 2.8 - Training programme (Patricia Mergen)
 1120 Overview of data management plans in WP5 (Carlos Rodriguez)
 1130 Discussion
- 1200 Lunch
- 1300 Split to tracks to work 2-3 hours on task implementation

Track 1 – Chair Wim Hugo, Rapporteur from CSIC/IBM (Task Leader)

- Architectural design – how to integrate with GCI, LifeWatch, and other networks

- EU BON portal
- Track 2 – Chair Mark Schildhauer, Rapporteur Éamonn Ó Tuama (Task Leader)
- Review and guidelines for using data standards
 - Design of monitoring sites – WP5
 - Data sharing tools – how to cover more data types and implement wider standards
- Track 3 – Chair Urmas Kõljalg, Rapporteur Patricia Mergen (Task Leader)
- Gap analysis and priorities for filling identified gaps in data coverage and quality – WP1
 - Registry and metadata catalogue
 - Helpdesk and training programme
- 1500 Coffee
- 1530 Report back and joint discussion, up to 45 minutes per track
- 1800 Trondheim and surrounding survey, by bus. Visiting interesting locations including Gaulosen Nature reserve. Binoculars are recommended.
http://www.google.no/search?q=gaulosen&tbm=isch&tbo=u&source=univ&sa=X&ei=826CUffbFlbptQbD_YGgBw&ved=0CDUQsAQ&biw=1920&bih=979
- 2000 Dinner at restaurant Aizuma, Kjøpmannsgate 53, hosted by EUBON

31 May - Friday: Where do we go from here?

- 0900 Timeline of actions
- T2.2 Standards review
 - T2.1 Architecture
 - T2.6 Helpdesk and training
- Pilot projects for 2015
- Shared data store infrastructure
 - EBV data flow
- 1030 Coffee break
- 1100 Group photo
 Travel details to airport
 Coordination with other work packages and external projects
 Upcoming events
 General discussion
- 1300 Adjourn and lunch in the restaurant of the Conference Centre, Høgskoleringen 9

Participants

1. Kessy Abarenkov, University of Tartu, Estonia; WP1
2. Donat Agosti, Plazi, Switzerland; WP3, WP1, WP2, WP8
3. Luis Aused Lopez, CSIC/IBM Global Business Services, Madrid, Spain; WP2
4. Mikel Díez Parra, CSIC/IBM Global Business Services, Madrid, Spain; WP2
5. Gregor Hagedorn, Natural History Museum, Berlin, Germany; WP2, WP9
6. Jonathan Giddy, Cardiff University, UK; BioVeL
7. Frank Hansen, Norwegian Institute for Nature Research, Trondheim
8. Wim Hugo, South Africa Environmental Observation Network, GEO BON; **ITF Member**
9. Matúš Kempa, Institute of Botany, Slovak Academy of Sciences, Bratislava; WP2
10. Urmas Kõljalg, University of Tartu, Estonia; WP1
11. Dimitrios Koureas, Natural History Museum, London, UK; WP3
12. Sjur Ringheim Lid, Institute of Marine Research, Trondheim
13. Wouter Los, LifeWatch, Netherlands; **ITF Member**
14. Patricia Mergen, Royal Museum of Central Africa, Tervuren, Belgium; WP2
15. Andreas Müller, Botanic Garden and Botanical Museum, Berlin-Dahlem, Germany; WP1
16. Stefano Nativi, Italian CNR, Florence, and GEOSS Infrastructure Implementation Board; **ITF Member**

17. Israel Peer, GlueCAD Ltd, Haifa, Israel; WP2
18. Johannes Peterseil, Umweltbundesamt GmbH, Vienna, Austria; **ITF Member**
19. Carlos Rodriguez, CSIC, Sevilla, Spain; WP2, WP5
20. Mark Schildhauer, National Center for Ecological Analysis and Synthesis, Santa Barbara, CA, USA; **ITF Member**
21. Éamonn Ó Tuama, GBIF Secretariat, Copenhagen, Denmark; WP2
22. Florian Wetzel, Natural History Museum, Berlin, Germany; WP1, WP2

Organisers:

23. Wouter Koch, Norwegian Biodiversity Information Centre, Trondheim; WP1, WP2
24. Nils Valland, Norwegian Biodiversity Information Centre, Trondheim; WP1, WP2
25. Hannu Saarenmaa, University of Eastern Finland, Joensuu; WP2

Other **ITF Members**, not attending this meeting physically:

- Gregoire Dubois, European Commission JRC
- Robert Guralnick, University of Colorado, USA
- Alex Hardisty, Cardiff University, UK
- Dave Thau, Google Inc., Mountain View, CA, USA

Annex II

EU BON Initial Informatics Workshop Trondheim, Norway, 29-31 May 2013

MINUTES

This is a markdown of the discussions as made by Hannu Saarenmaa (HS)

DAY 1

For presentations, see the slides.

During his presentation Wouter Los (WL) asked whether there will be an ICT reference model for EU BON? We need to clarify the roles of the LifeWatch service centre vs EU BON helpdesk. LifeWatch international coordinating committee has strong overlap with EU BON - avoid overlap. EUDAT data infrastructure could potentially be backbone for data storage of EU BON. Data upload must go to simple stores. Use data replication for safety. Let's organise small technical working meeting LifeWatch, EU BON and other projects to consider integration of the various contributions.

DAY 2

T2.1 Carlos Rodriguez (CR) presented. SOA, brokering. Data integration through a central metadata repository. Links to full data, but not aggregation. Main capabilities: Registry, metadata catalogue, portal, semantic mediation, thesaurus, component-oriented infra. Workplan to define the architecture, component interfaces, execution components.

Stefano Nativi (SN): It is good to continue with CGI. If need access to climate change data, CGI can provide that from the GEOSS data core. Provide open and free data.

Mark Schildhauer (MS): This plan is still at high level. Experience of DataONE is that building the system requires a huge amount of effort. A dozen sw engineers have been working for 3 years. Authentication and site based solutions are complicated. Leverage on existing efforts is instead of building from start.

Wim Hugo (WH): The background operations layer is going to require much effort.

CR response: We have clear stakeholders in WP5 and will use that information as reference how to design the architecture.

T2.2 Eamonn O Tuama (EOT) presented. The milestone report has gone out in time.

MS: Do not only limit to Europe but look globally. Discerning trends requires monitoring power. Are there modelling experts in the task?

T2.3 Patricia Mergen (PM) presented.

MS: What is meant by digitisation?

PM response: It is databasing. Good modelling requires data from the same periods. Such data is still often in paper form and needs to be digitised.

T2.4 EOT presented

SN: The metadata catalog should be accessible from GCI. It supports many protocols.

Frank Hanssen (FH) asked: Is there any role for EUDAT?

T1.1 Urmas Koljalg (UK) presented. The task is looking for partners for evaluation of data sources of remote sensing.

Nils Valland (NV): Norwegian space centre has experience. We may reach out to them as associate partner.

Israel Peer (IP): Will gap analysis be aligned to WP4?

NV: Are you considering habitats? For instance EEA has such data.

T2.5 CR presented

Johannes Peterseil (JP): There is direct link to LTER. What is the the definition of a site? Is it the same as LTER site? Is Donana a site?

CR response: Work is going on, no answer yet.

JP: How you document a site?

CR: Infobase is very complicated. We are trying to do something simpler.

JP comment: How is site described need to be discussed.

CR: We need tools in portal by which documenting is easy.

T2.6 PM presented the helpdesk services and documentation.

WH: Will there be references to documentation of standards? Something to consider. Link to standards.

SN: OGC vs Restful was not clear. Do not use the term "OGC model" - it is not a model.

T2.8 PM presented training programme.

MS: Rob Guralnick and Walter Jetz operate week-long BI trainings. DataOne trainings are available.

PM response: We will keep lists of other trainings. We also collaborate with JRS Foundation who are doing huge trainings. We try to combine efforts.

T5.1 CR presented

MS: Is an EU BON site interested in all aspects of sites?

CR response: Today scientists come and go with their data. Our new Director wants data also stored in the site directory.

=====

GROUP WORK REPORT BACK

Track 1: See Luis Aused (LA) presentation.

The goal of Task 1.1 is to build a roadmap and implementation plan by January 2014. It will describe who will do what and when, and describe involvement of partners. There will be a technical advisory group. The architecture will describe the components, including an enterprise service bus, and a number of services.

MS: You will get all sorts of heterogeneous data. How do you deal with that? Into database?

LA response: We have a reference data model. You can do some searches to the metadata.

MS: Another question is the data itself. There will be some data and there will be documents, pdfs. Rows and columns can be in any order. Can this be standardised somehow?

WH: Our experience is that we go to scientists and give them guidance: organise your data this way, but they do not often listen to us. However, mapping may be possible to common data model from these poorly organised data.

WL: Who will do this mapping will remain a challenge. Automatic annotation is a possibility.

LA: We cannot have all the data in portal, but metadata we can organise.

JP: Community of data providers will be diverse. Data can be reported to multiple communities. Virtual translation must be done with the involvement of the data owner. A data cache is needed. Basket of data is needed for small sites and users that cannot setup their own data provider. That is an important use case.

SN: Push standardisation as much as possible. But it is not going to solve all aspects. You need also support the mapping approach.

WH: EU BON or GEO BON could host the mappings.

LA: We do not take the analogy of shopping list.

Gregor Hagedorn (GH): For the download part, look at SECAL open source system for distributed datasets. The NHM in London is implementing it for all its biodiversity data.

Track 2: See EOT presentation.

Donat Agosti (DA), GH: LSIDs are dead. URIs are being promoted by CETAF.

WH: Chemical data standards will be gotten from external communities.

Jon Giddy (JG): Workflows of R-scripts can be accessed through standardised workflows.

EOT: Who would develop workflows in EU BON?

DA: Licensing. We need to look if we really want standardise something like CC license. Machine processable licences is what we need.

WL: That is right. We cannot click 100 000 times yes or no. We needs to standardise the licenses.

WH: There is work in progress in legal interoperability. Being discussed in CODATA.

EOT: GBIF is watching this.

MS: DataONE has some standardisation of real data.

SN: Data sharing group of GEOSS is working on this.

GH: You cannot control something w cc which is not copyrightable. You cannot use (c) to hide protected species. CC licenses are copyright-based.

PM: We should be very careful to keep copyright and citation. Once Europeana moved to CC0, many data providers left. GBIF is not doing this which is good. Do not follow them. If you take CC0, anyone can use it without citing

GH: Citing is important for scientists. It is not about copyright but norms in science. Concern is valid. Copyright is not the way to solve it. W3C ... may.

GH: In EU we have de novo right which is relevant to databases. It was created for phonebooks. It is not copyright, but if someone creates a database of taxonomy, nobody else may... CC0 is very good if you don't want to enforce copyright. Make citing easy technically.

Track 3: See PM presentation.

MS: Many of these activities could proceed without waiting for infra.

WL: Many like Citizen science communities have their own training programmes. How to join them?

PM response: We are already thinking of joining efforts, but the experience is that it will require lot of adjustments and may compromise the course. Be careful.

WH: Gap analysis could also include temporal gaps.

=====

DAY 3

Actions on Tasks:

T2.2 EOT will send templates in 2 weeks. Be concise in your writing and send word doc back by 1 September. Review of a combined document will then follow by 1 November.

T2.1 CR said that first document version will be available for review in September. The second version December.

HS: A service catalogue will be needed. It can be gotten from BioVeL.

WH: There will be constraints for the architecture. Please be clear on them.

CR: We will need datasets for testing.

PM: Gregoire Dubois has test implementaions, modelling services,

NV: NBIC has redlisted species data
EOT: GBIF has prototypes of for new WMS, WFS.

T2.6 PM said that basic informatics training is needed as service to scientists. Shall EU BON implement that?

MS: Indeed NCEAS has lots of applicants for such training. Training on basic data management is needed.

PM will now be doing one month of training for African visitors. We have a training centre but can training be only done in one place?

Pilot projects:

HS presented an 1) architecture pilot for a new open-access platform and 2) a use case pilot for Essential Biodiversity Variables.

He began by showing important 4 slides from WL on “We need to change the way we deal with data.” Users still generate data via standalone tools to standalone storage. Siloed approach today. We need to move from piggybank approach to a databank infrastructure. Data need to move automatically like money transfers. Databanks with credit - build the trust.

1) Open access data sharing platform

GH: Support working with DataONE on the data sharing platform. We need to think of the maintenance costs of such platform after the EU BON project. It will require continuous updates. Open source is the only possibility. We do not just want to take up DataONE solutions but contribute to them as well.

MS: DataONE will probably be funded for another 5 years. The new contract includes developing a sustainability model after the NSF funding is over. This could be linked with NSF requiring use of this platform for data storage in its grants. International partners can also help in sustainability. DataONE wants to expand its partnerships. For instance GEOSS- DataONE partnership is on table. EU BON contributions would fit well in this.

NV: Is EUDAT an alternative?

GH: EUDAT is a consortium of large data centres. It does not include such services as DataONE has for storage. EUDAT is similar to IRODS in the USA.

MS: DataONE is using IRODS.

GH: EUDAT is agnostic about the interoperability solution. DataONE technology can be used together with EUDAT.

NV: Does DataONE meet needs of users about scientific credit for data sharing?

MS: NSF requires a data management plan from its grantees. The synthesis centres have shown that data shoring is a valid approach.

NV: Data theft is still a concern to many which is only slowly going away.

GH: We are not fully addressing this problem. EU BON needs to meet this challenge. Together with DataONE we can do it. We can build an European contribution to DataONE. We have only two options: DO it from scratch or use existing solutions from DataONE.

NV: Will this solution help releasing the data?

WH: Broadly agree, but many potential providers are not in position to use these solutions. Using generic infrastructure is still too hard. This is only infrastructure of the "last resort" if your institution does not offer any. Data citation will be required. We will have only little time to do all this.

MS: In molecular biology it already works.

SN: We hope at GEOSS that big data centres will be growing, but we also have to facilitate many diverse solutions. For instance in the last plenary we discussed role of private companies. Google etc increasingly offer free services. These are appealing to citizen science networks who may not want to come to common data centres.

Sjur Ringheim Lid (SRL): Where will the data go after 5 years when EU BON or DataONE is over? Scientists will not trust their data to them, if there is no plan.

MS: Libraries are increasingly becoming data custodians.

HS: In Finland there is a national digital library that is building a long-term persistent data storage with a perspective of hundreds of years. National museums of natural history have similar perspective and we can rely on them in the long run.

GH: EUDAT will be charging research projects for their stored data. But who is paying in the long term for the projects? Libraries and natural history museums guarantee the longevity. These are memory institutions and it is their core business.

DA: Helpdesk could offer training on the principles and benefits of open access to data.

PM: Yes, we could do that together with the OpenUp! project.

NV: I like this in principle. The partners could organise conferences of open access at national level, and introduce national implementations.

IP: There is still the question how to start piloting in practice.

HS: The willing partners can start piloting by installing DataONE member nodes. These could be populated with data from the networks they work with. We would then share experiences, and perhaps later build a coordinating node.

2) EBV pilot

HS: This could be done as GEOSS AIP-7.

SN: The call for API-7 will be opened in February 2014 and end in September 2014. Then a movie for the GEOSS Plenary needs to be prepared.

HS presented the agreements on the pilot projects.

Florian Wetzel (FW): Will be working on evaluation of data sources. Will be contacting partners for advise on the data sources.

Events:

PM: WP2 members should remember to inform the coordinator on relevant events.

PM: EU BON abstract for TDWG conference would be needed. WP leader and task leaders will organise this.