Meeting Report

8th European GBIF Nodes Meeting
Lisbon, 19-21 April 2016

Workshop on sample-based data publication
Lisbon, 18-19 April 2016

nó português do GBIF

Index

Acknowledgements

1. Introduction

The 8th European GBIF Nodes Meeting took place in Lisbon between the 19th and 21st of April 2016. The Venue of the event, which was organised by GBIF Portugal, was Instituto Superior de Agronomia's (GBIF Portugal's host institution) Salão Nobre.

The extended Programme of the Meeting, which began on the 18th of April and ended on the 22nd of April, included a Workshop, an Ice-breaker, an Official Dinner and an Excursion (Table I)(see Annex IV for a selection of pictures of the different events).

The agenda of the Nodes Meeting and the Workshop on Sample-based Data Publication are included at Annexes I and II, respectively. The list of participants in the meeting and in the workshop is presented in the Annex V.

This report compiles all materials produced/used in the meeting and in the workshop. It is also included, in Chapter 4, the report of the evaluation assessment both events that was performed online after the meeting. Finally, it includes the indication of web/social media/press releases performed to promote the events in Portugal and globally,

	Mon, 18 Apr	Tue,19 Apr	Wed, 20 Apr	Thu, 21 Apr	Fri, 22 Apr
09:00h - 13:00h	Workshop	Workshop	Nodes Meeting	Nodes Meeting	Excursion
14:00h - 18:00h	Workshop	Nodes Meeting	Nodes Meeting	Nodes Meeting	Excursion
19:00h	Ice Breaker		Official Dinner		

Table I - Extended Programme schedule

2. 8th European Nodes Meeting - Summary and action points

The main goal of the 8th European Nodes Meeting was to advance the contribution of the nodes for the European Work plan 2015-2016, and simultaneously promote the coordination and contribution of the GBIF nodes of Europe to the global work programme. The various topics build on the areas for which nodes have been increasing their capacity, or have a raised interest, which were:

- Nodes Portal

- Sample Based Data

- Uses of Data, user-science-policy interface

- Data quality, validation (and fitness for use)

Furthermore, the meeting was the opportunity to define goals, priorities and actions for the European Work programme 2016-2017, build on the actions points previously set for 2015-2016. Finally, the meeting resulted in an excellent opportunity to exchange experiences between the different nodes, and get and update on the activities of the Secretariat developed on the Work programme and on the Strategic Plan 2017-2021.

2.1. Main results of the meeting

2.1.1 . Preparatory Survey

The meeting was preceded by a preparatory survey, which results' report is available at the Community site. This indicates many and diverse activities and projects on which the European Nodes are involved. The activities with more nodes involved are, by order of number of nodes:

- Data mobilization/publishing (8)

- Data Portal (8)

- Invasive Alien Species (3)

- Capacity Enhancement (2)

Several other topics were mentioned by the nodes. The report also summarizes the items identifies by the nodes in a SWOT analysis several items, which main topic for each of the entries, with the respective number of answers, were:

- Strength: good networking (8)

- Weakness: lack of resources (7)

- Opportunities: new data portal/research infrastructure (3)

- Threats: insecure funding (10)

The summary of the survey also identifies main collaborations between European Nodes, main items on expertise, knowledge exchange, tools, data and data management were nodes expressed installed capacity, and main areas for collaboration at the regional level.

2.1.2 . Thematic areas

Four thematic areas were discussed in break-out group sessions, as an opportunity to share best practices, which were reported back to the plenary: Main outcomes were, per topic:

Nodes Portal

- CoopBioPlat and Encounter Bay as two projects implementing and documenting the tool

- Opportunity offered by EGI for hosting in Europe

- Sharing of the backend

- Finding a place to share the tools globally

- Organizing further regional meetings while Dave Martin is still in his current position

Full notes and outcome available.

Sample based data publication

The discussion on this group summarised the outcomes of the pre-meeting workshop on Sample-based data publication, and contributed to a series of recommendations in this document available, on several topics documentation, standards, visualisation and mobilisation.

Uses of data, users-science-policy interface

- Nodes are closer to the data providing end of things, so it is easier for them to identify use cases for providers rather than users.

- There is more and more organizations that link data from GBIF as part of their own services.

- There's a component of planning and engaging both users and publishers, not only to record use that has already happened.

- The IAS is a very good case to start with.

Full notes and outcome available.

Data quality, validation (and fitness for use)

Several action points resulted from the discussion on this topic:

- Data validation toolkit (GBIF.no) – actions: Start a GitHub project for sharing source code and collecting feedback; Install and dedicate a demo IPT for uploading test data
- Data fitness for use task groups: reference to the reports by the the GBIF Task Groups on Data Fitness for Use in Agrobiodiversity and Data Fitness for Use in Distribution Modelling, and TDWG Data Quality Interest Group
- Georeferencing actions - Mobilize data for coordinateUncertainty; Respond to AnnoSys annotations

2.1.3 . European Workplan 2016-2017 - Summary of action points

The whole group reviewed and discussed several activities that contribute to the European

Workplan 2016-2017, organised by action points firstly established in the 7th European Nodes Meeting. The action points, set by main strategic priority – deliver relevant data; improve data quality; fill data gaps; enhance biodiversity information infrastructure; empower global network – were reviewed and updated, and also attributed to responsible persons for the respective action points. For each of the points, a contact for the Secretariat was also indicated. The summary and detailed identification of the action points is available in this document.

2.1.4 . Other topics of the agenda

Ten short presentations by several nodes reported on collaborations, nodes activities and collaborations. Special focus on collaboration opportunities, including the follow up and resubmission of the COST Action.

It was recommended to the Nodes to perform the Capacity Self Assessment exercise to assess their national/institution capacity strength and needs.

Two presentations by representatives from Russian institutions allowed participants to follow the advances in GBIF activities in Russia, as well as in biodiversity information management.

3. Workshop on sample-based data publication – summary and recommendations

The workshop on Sample based data publication was organised as a side event of the 8 th European Nodes Meeting, by Dag Endresen (GBIF Norway), Kyle Braak (GBIF Secretariat), Christian Svindseth (GBIF Norway) and Anders G. Finstad (GBIF Science Committee, NTNU University Museum in Trondheim).

The topic of the workshop was identified has being of high interest for the generality of nodes, as a follow up of the first training on sample based dataset publication performed in the 13th Global Nodes Meeting. However, the goals of the workshop were not focused on training, but in:

- discussing specific aspects and issues with the publishing of sample-based, event-core data and datasets;
- identifying issues and suggesting possible improvements for this new data type;
- exploring issues and solutions, using IPT or other approaches.

The agenda of the workshop (Annex II) was organised in two parts, the first one for short explanation of concepts, examples and recent advances on the topic, and the second part organised in break-out groups to analyse and discuss specific dataset examples. As of the first part, some general questions guided the presentations and discussions:

- What is sample-based data?

- How do we recognize if a dataset is sample-based?

- Can the same dataset be represented using different core types?

- How to present and visualize sample-based data in the GBIF portal?

- How do we best model sample-based data for reuse in new contexts?

- How to design sample-based data APIs and web services.

As a preparatory task for the workshop, participants were invited to bring their sample datasets, fully documented, to be prepared for publication as sample based dataset. Groups were organised around the several sample datasets, in order to resolve several questions related to the publication process. The group should report on these outcomes

- Were you able to find out enough information (metadata) about the dataset to understand it? Could you fill in the missing columns in the Overview - Example Datasets spreadsheet?

- Which issues (topics below) did you encounter while working on the example dataset?

  - What format (event vs. occurrence) to use, also which extensions

  - How to handle presence / absence data

  - How to record the sampling protocol, abundance, other DwC fields

  - How to transform the raw data into DwC tabular form

- For each issue you encountered, did you find a solution / recommendation? Please describe.

- Did you successfully publish the example dataset into the sample event format?

- What documentation did you use to learn more about how to publish sample event data? E.g. what DwC terms to populate, how to populate DwC terms, list of extensions available for use, etc.

- Was the documentation sufficient, or was there something missing? E.g do you need this documentation in your local language, or is English sufficient?The full detail on the five datasets, including summary of the outcomes of the group work, and post-workshop follow up, are included in Annex II, and also on the agenda [available online](#).

From the workshop discussions, a set of recommendations highlighting missing features/functionalities/tools was made for future developments on sample-based dataset publication. The full report is included in Annex III, and also [available online](#), but some highlight can be made in four main areas:

Documentation:

- the GBIF/EUBON document IPT Sample Data Primer needs to be updated with a series of improvements identified in detail in the recommendations document;

- a series of suggestions was made to improve the GBIF document IPT How To Publish Guide

- it was recommended that authors from GBIF and the wider biodiversity community should publish a scientific article explaining how sampling event data can now be standardised and published through GBIF.org.

Standards

Several changes to DwC and the current set of extensions and vocabularies has been proposed, based on feedback collected during the workshop, concerning the addition of new terms and controlled vocabularies. Other recommendations concern the development of a standard cross tab table template, the evaluation of the use of the GBIF Relevé Extension and an investigation on how to properly record presence/absence data.

Visualisation

Concerning visualisation of sample-based data, and based on several existing systems, aseries of issues were submitted to the GBIF issue management system.

Additionally, participants helped collect the following set of bugs and enhancements submitted to the GBIF issue management site.

Mobilisation

In order to streamline sample-based data mobilisation, and identify new data sources, the involvement of the nodes' community is essential by:

- approaching thematic groups or EU funded projects doing standardised sampling such as BioFresh in order to help mobilise sampling event datasets;

- assistance in identifying sampling protocols that are in common use, with the aim of producing a controlled vocabulary for sampling protocols.

- identifying users of Turboveg software within their network, and educate them it is possible to export vegetation data in DwC-A format using Turboveg's built-in DwC-A export feature

Another recommendation is the republishing of occurrence datasets as sample-based datasets

4. Presentations, material, multimedia presented at the meeting

Here is a list of links to the multimedia support materials used by each speaker during his/her oral communication, listed in chronological order according to the Meeting/ Workshop Programme (see Annexes I and II):

Communication/Materials Title	Speaker/Author	Link
8th GBIF European Nodes Meeting Preparatory Survey	Anne-Sophie Archambeau	http://www.gbif.pt/sites/default/files/Survey.pptx
Nodes Committee and Regional Work Plans & activities	Hanna Koivula	http://www.gbif.pt/sites/default/files/NC_and_Regional_WP.pptx
GBIF Strategic Plan 2017-2021	Alberto González-Talaván	http://www.gbif.pt/sites/default/files/20160415%20-%20GBIF%20Strategic%20Plan%202017-2021%20-%20small_0.pptx
GBIF Progress Update - April 2016	Kyle Braak & Alberto González-Talaván	http://www.gbif.pt/sites/default/files/2016-04%20-%20GBIF%20Update%20Slides%20For%20Regions%20-%20small_0.pptx
EUBON Update	Anders Telenius	http://www.gbif.pt/sites/default/files/EU%20BON%20LifeWatch%20ENM%202016.pptx
LifeWatch Update	Anders Telenius & Rui Figueira	http://www.gbif.pt/sites/default/files/EU%20BON%20LifeWatch%20ENM%202016.pptx
COST Action and European GBIF Nodes	Anne-Sophie Archambeau	http://www.gbif.pt/sites/default/files/COST-Presentation.pptx
CoopBioPlat - A cooperative framework for building a common platform to serve biodiversity information at national level	Cristina Villaverde	http://www.gbif.pt/sites/default/files/CoopBioPlat-CV.pptx
GBIF_PALOP - Promote GBIF in the African Portuguese Speaking Countries Through Documentation and Seminars	Rui Figueira	http://www.gbif.pt/sites/default/files/GBIFPALOP_RuiFigueira.pdf
Citizen Science and species occurrence data in Europe\n\nCitizen Science Species Occurence Data Availability Europe Spreadsheet	Nils Valland	http://www.gbif.pt/sites/default/files/European_CitizenScience_compiledNBIC_20160418_0.pptx\n\nhttp://www.gbif.pt/sites/default/files/CitizenScience_SpeciesOccurence_DataAvailabilityEurope_20160219_NilsValland%40NBIC_0.xlsx
A Generic Annotation System - Integration into the GBIF Data Portal	Jörg Holetschek / Lutz Suhrbier & Wolf-Henning Kusber	http://www.gbif.pt/sites/default/files/AnnoSysIntegrationInGBIFdataPortal.pptx

Best practice exchange: Nodes Portals thematic group-session Notes (Materials for discussion)	n.a.	https://docs.google.com/document/d/1UuvXbfrU73MurfVntWgqihDbMi4jcYWYEIcAxts-ZBg/edit?usp=sharing (http://community.gbif.org/pg/pages/view/49268/work-plan-group-green-sharing-and-jointly-developing-nodes-portal-solutions)
Best practice exchange: Sample-based data (Materials for discussion: Pre-meeting workshop on sample-based data)	n.a.	http://community.gbif.org/pg/groups/47949/samplebased-data-publishing-interest-group/ (https://goo.gl/oqg2OR)
Best practice exchange: Uses of data, users-science-policy interface, reaching out to the cultural domain_Session Notes (Materials for discussion)	n.a.	https://drive.google.com/open?id=15D5qwWQcnY0oirpx8kIS11iTqCuxCCuq-CvI1qEvg0U (http://community.gbif.org/pg/pages/view/49063/ http://community.gbif.org/pg/pages/view/49334/work-plan-group-red-improving-relevance-of-gbif-for-sciencepolicy-initiatives https://docs.google.com/document/d/1AZ11Axu747H7e386R6KaIC4FjWoSbdjn3V4kSzAp3oE/edit?usp=sharing)
Best practice exchange: Data quality, validation (and fitness for use)-Session Notes (Materials for discussion)	n.a.	https://docs.google.com/document/d/197Btzw4CXgQl6d6Bjcsb63wNbXQ2yoTO_HHvKZfRjmo/edit (https://docs.google.com/document/d/197Btzw4CXgQl6d6Bjcsb63wNbXQ2yoTO_HHvKZfRjmo/edit ; http://community.gbif.org/pg/pages/view/49262/work-plan-group-blue-improving-the-quality-of-gbif-mediated-data https://github.com/GBIF-Europe
OpenUp! - The Natural History aggregator for Europeana, Europe's digital library	Jörg Holetschek	https://gbif.isa.ulisboa.pt/sites/default/files/OpenUp.pptx
Discussion around the next Action plan (COST): Biodiversity informatics curriculum-Session Notes (Materials for discussion)	n.a.	https://docs.google.com/document/d/1QA6B3EJDrK0k2Fbqyb2cnu7188VIQ7r8xxoF_39E7YE/edit (http://community.gbif.org/pg/pages/view/48692/breakout-group-towards-a-curriculum-for-biodiversity-informatics)
Discussion around the next Action plan (COST): Digitization and repatriation (Materials for discussion)	Anders Telenius / Alex Asase & Anders Telenius	http://www.gbif.pt/sites/default/files/GNM%2013%20B-O%20Yellow%20Digitization%20and%20Repatriation.pptx (http://community.gbif.org/pg/pages/view/48970/

		breakout-group-common-approaches-to-digitization-and-repatriation)
European Nodes Work plan 2016-2017: Defining goals, Priorities and Actions for the coming 12 months-Summary of actions points	n.a.	https://docs.google.com/document/d/13V4Ha1NBoCYyJWGx6zjkHX_CHispXT0hCxfL5TVc0fM/edit
GBIF activities in Russia	Maxim Shashkov / Maxim Shashkov & Natalya Ivanova	http://www.gbif.pt/sites/default/files/Shashkov_Ivanova_MSh_NI.pdf
Zoological Institute of the Russian Academy of Sciences — first steps in global data publishing integration	Roman Khalikov	http://www.gbif.pt/sites/default/files/Khalikov_GBIF.pptx
European Nodes Work plan 2016-2017_Wrap up and Conclusion (Other Resources mentioned: Opening the book on open access Capacity self-assessment guidelines for national biodiversity information facilities)	n.a.	(http://www.openaccess.nl/en/events/opening-the-book-on-open-access http://www.gbif.org/resource/82277)

5. Evaluations of the Meeting and Pre-Meeting Workshop

After the Meeting the participants of both the meeting and the pre-meeting workshop were asked by e-mail to submit their evaluation of the events by filling-in an online evaluation form (see Annex VII) at meeting website (http://www.gbif.pt/EuropeanNodesMeeting).

5.1. Meeting Evaluation Survey Results

Of a total of 22 participants (24 minus 2 organiser) 16 (73%) participated in the online Evaluation survey, although some did not answer some questions.

5.1.1 . General characteristics

Most of the respondents reported to be either satisfied or very satisfied with most of the general characteristics of the Meeting (e.g.: Information Sharing, Interaction with Peers and opportunities to voice own opinions). Aspects that respondents were, overall, less pleased with, with a few reporting to be Dissatisfied, were Setting of the meeting agenda, Answering of doubts and Usefulness for future work.
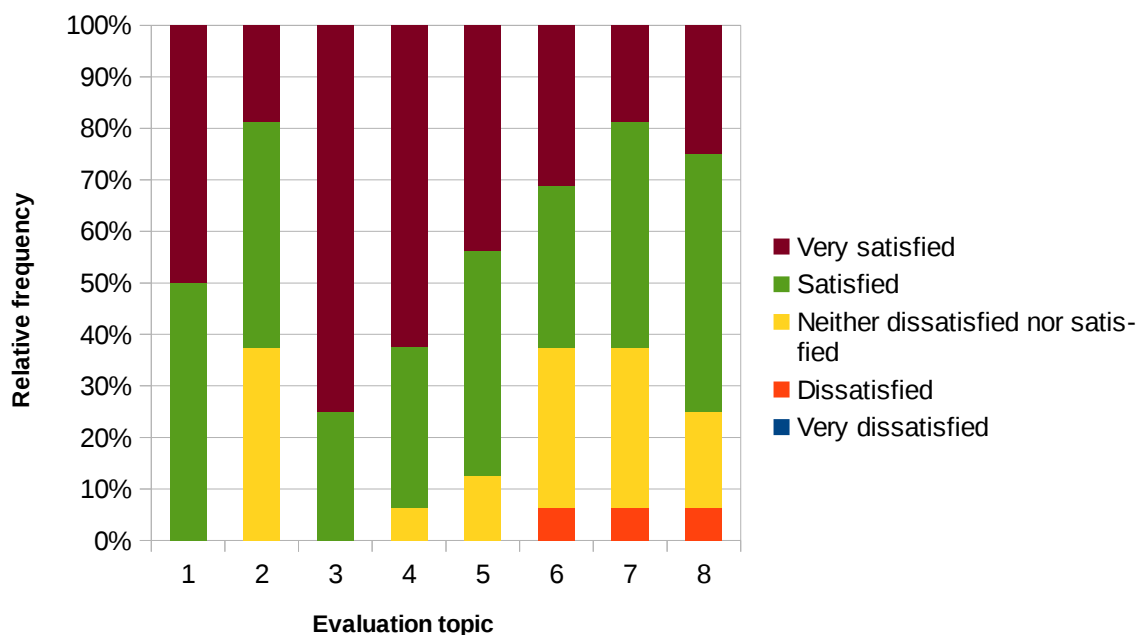

Fig. 1 - Participants' satisfaction levels regarding Meeting general characteristics (*1 – Information sharing, 2 – Setting common priorities, 3 – Interaction with peers, 4 – Opportunities to voice opinions, 5 – Strengthening/establishing collaborations, 6 – Setting of the meeting agenda, 7 – Answering of doubts, 8 – Usefulness for future work*)

5.1.2 . Venue and Organisation

Similarly, satisfaction levels regarding most of the aspects assessed were high, with 56-100% of the respondents reporting to be "Very satisfied" about 7 of the 9 topics analysed. The "Excursion" and "Helpfulness of the organisers" scored the highest (100% of the meeting delegates that filled-in the Meeting Evaluation form were "Very Satisfied" about both), while "Logistics and Practical organisation", "Meeting Participant packet", the "Hosted Dinner" and the "Ice-breaker" only registered positive assessments (all respondents said they were either "Very Satisfied" or "Satisfied").

The two topics on which participants satisfaction did not rate so high were "Adequacy of the Venue" and "Suitability of the furniture and equipment". In the first case, about 43,8% of the respondents said they were "Neither dissatisfied nor satisfied" while, in the second case, "Dissatisfied" represented 18,8% of the answers gathered. The one aspect referred in the final comment section of the survey about these matters was the poor acoustics of the room: one participant said "it was often difficult to understand other participants", while another reported "big difficulties to hear people in the room, even when using the microphones" and, a third one pointed "The venue was otherwise fabulous, but the acoustics was so bad, that only small breakout group discussion was fruitful. Plenary session discussion was difficult with microphones".
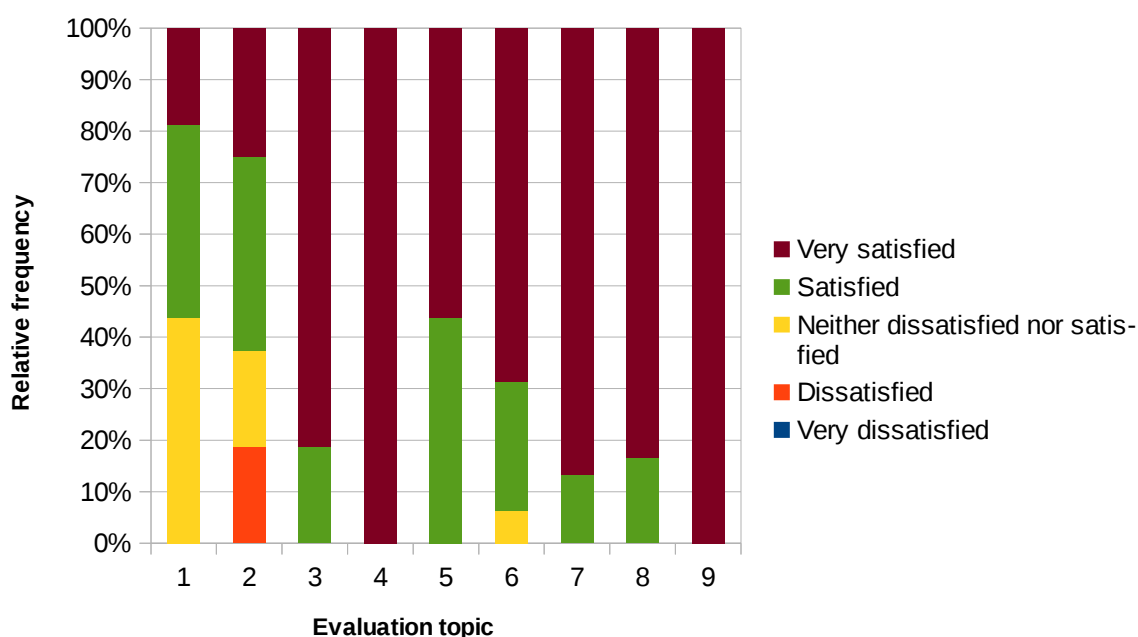

Fig. 2 - Participants' satisfaction levels regarding Meeting general characteristics (*1 – Adequacy of the Venue, 2 – Furniture and equipment, 3 – Logistics and practical organisation, 4 – Helpfulness of the organisers, 5 – Meeting participant package, 6 – Lunch and coffee breaks, 7 – Hosted dinner, 8 – Ice-breaker, 9 – Excursion*)

5.1.3 . Objectives

Over half of the respondents considered that each of the meeting objectives was either "Fully attained" or "Attained". Nevertheless, 28,6-42,9%, depending on the objective, were of the opinion that it was "Neither attained nor not attained". "European Nodes contribution to GBIF's Strategic Plan 2017-21" and "Updating from the Secretariat" were the only objectives to have negative assessments with 6,6% and 14,3% of the respondents reporting that it was "Not attained" and "Not attained at all", respectively.
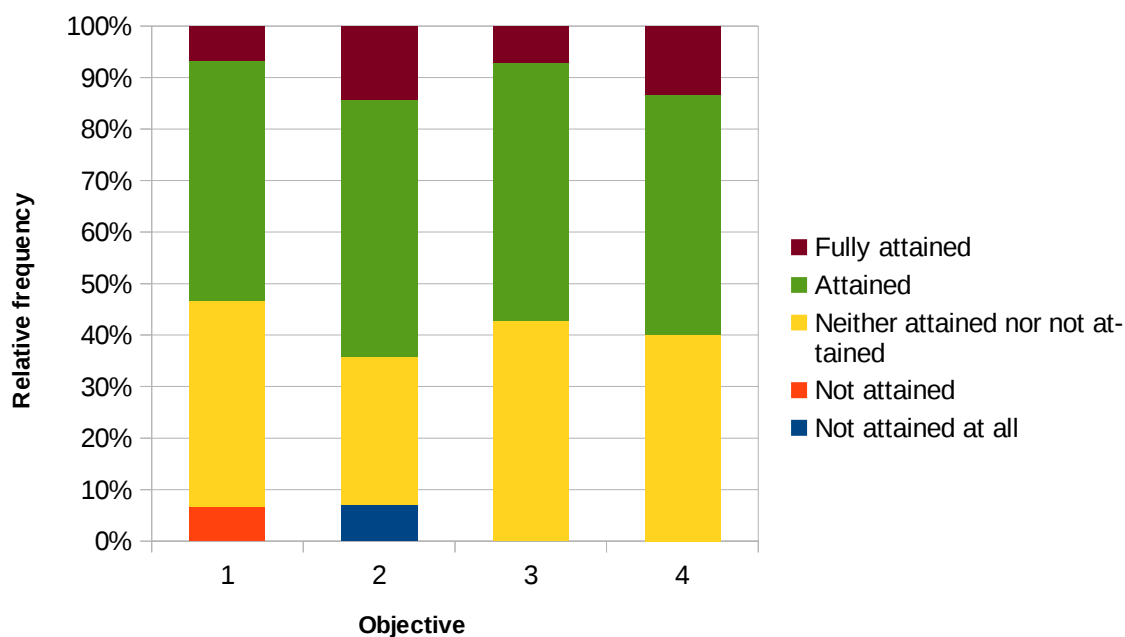

Fig. 3 - Participants' evaluation regarding the attainment of the Meeting objectives (*1 - Contribution to GBIF's Strategic Plan 2017-2021, 2 – Updating by the Secretariat, 3 - Updating/Reporting by European Nodes, 4 - Shaping European Nodes Work plan 2016-2017*)

5.1.4 . Parallel Session I

The survey respondents that participated in the "Sample-based data publication" group (N=8) were pleased with the work session, not having assessed "Negatively" or "Very negatively" any of the session aspects. "Engagement in the Discussion" scored higher with only "Positive" and "Very positive" assessments. At the other extreme, "Outcomes of the session" had the highest "Neither negative nor positive" assessments – 37,5% (N=3), with "Very positive" and "Positive" assessments accounting for 25% and 37,5% of the respondents answers, respectively.
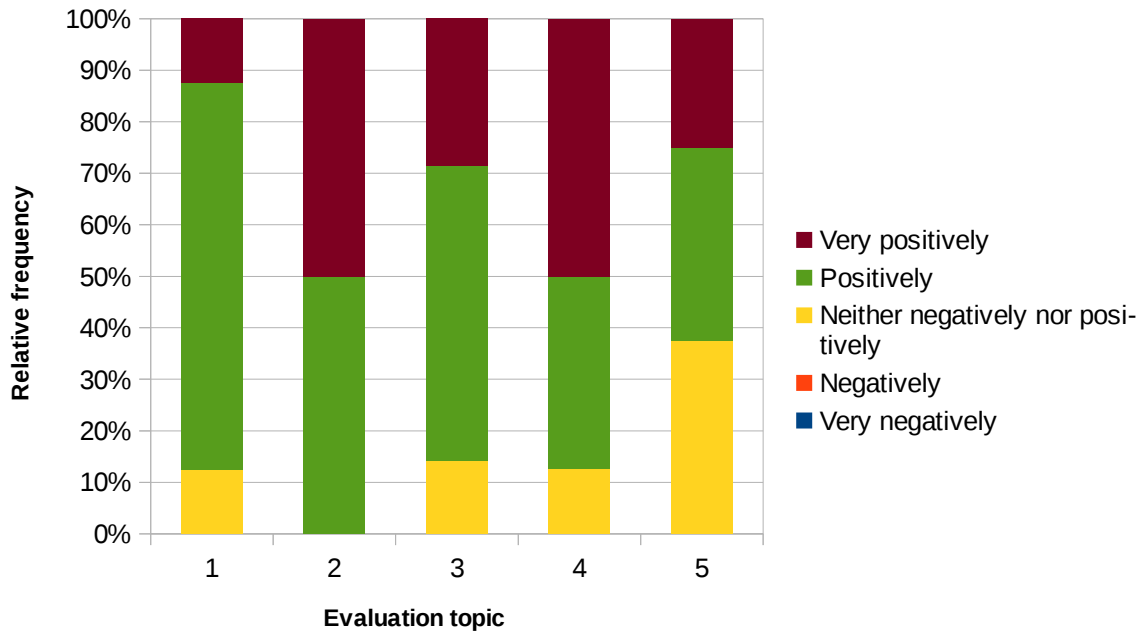
Fig. 4 - "Sample-based data publication" group participants evaluation of the work session (*1 – Preparation in advance, 2 - Engagement in the discussion, 3 – Session facilitation, 4 – Depth of the discussion, 5 – Outcomes of the discussion*)

Like in the other break-out group, there were no "Negative" or "Very negative" assessments of the "Nodes portals (and other tools)" group work session. However, neutral assessments accounted for over 60% (N=5 of a total N=8) of the answers regarding "Preparation in advance" and 50% regarding the "Depth of the discussion". "Session facilitation" was the most favourably assessed aspect with 85,7% (N= 6) of the respondents (N=7) rating it "Positively" or "Very Positively". 75% of the respondents assessed the "Outcomes of the discussion" either "Positively" or "Very positively".
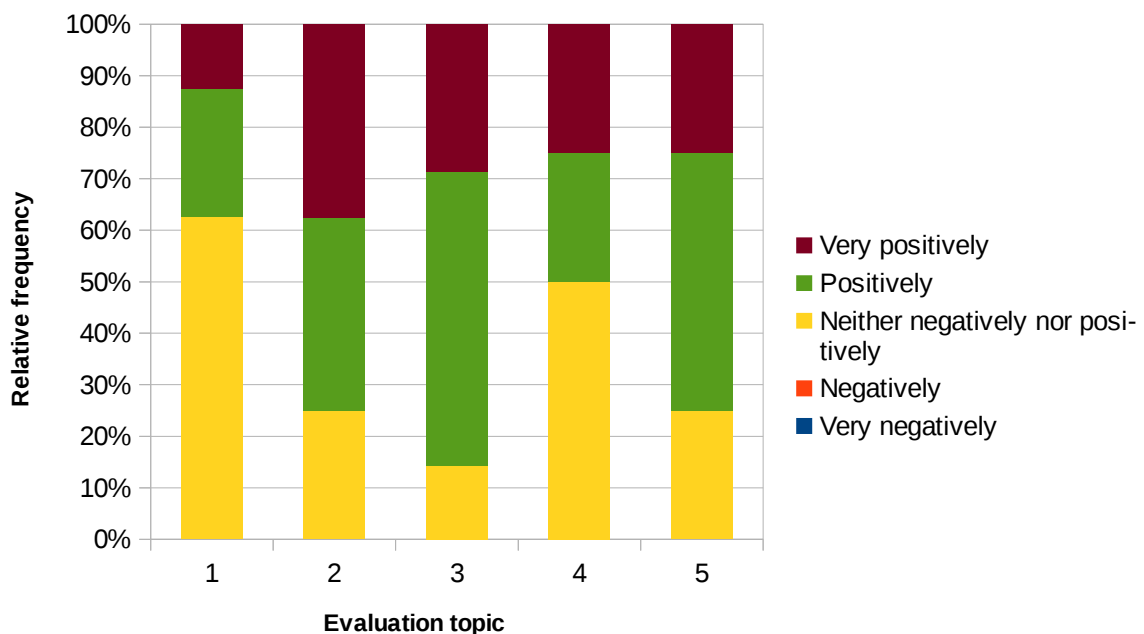

Fig. 5 - "Nodes portals (and other tools)" group participants' evaluation of the work session (*1 – Preparation in advance, 2 - Engagement in the discussion, 3 – Session facilitation, 4 – Depth of the discussion, 5 – Outcomes of the discussion*)

5.1.5 . Parallel Session II

Assessments of the "Data quality, validation" group work session (N=8) were not as good as those of the previous group work sessions. In fact, "Very positive" assessments accounted for, at the most, 12,5% (N=1 for any of the work session's individual aspects evaluation, except for "Outcomes of the discussion", which was assessed "Positively" by only 25% (N=2) of the respondents and "Very positively" by none. Nevertheless, there were no "Negatively" or "Very Negatively" assessed aspects and three of the five items ("Engagement in the discussion", "Session facilitation" and "Depth of the discussion") were assessed "Positively" or "Very positively" by, at least, 50% of the respondents.
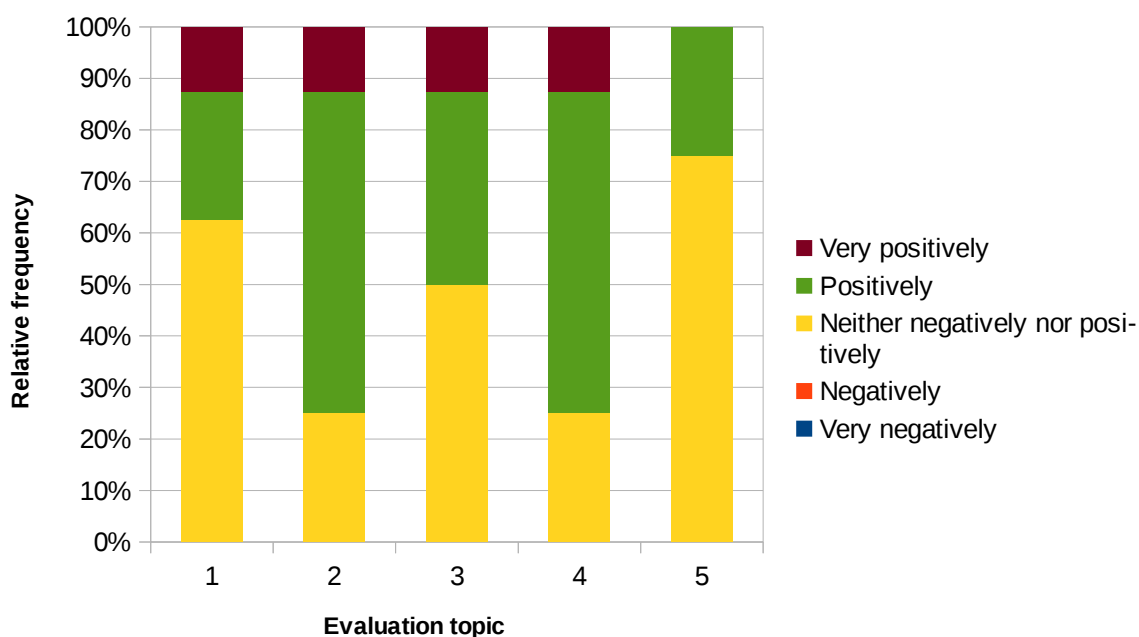

Fig. 6 - "Data quality, validation" group participants' evaluation of the work session (*1 – Preparation in advance, 2 - Engagement in the discussion, 3 – Session facilitation, 4 – Depth of the discussion, 5 – Outcomes of the discussion*)

The "Uses of data, user-science-policy interface and reaching out to the cultural domain" group work session was the first one with "Negative" and "Very negative" assessments, which accounted for 25-37,5% of the ratings (N=2-3 of a total N=8), depending on the item. Furthermore, the only aspect to have at least 50% "Positive" or "Very positive" assessments was "Engagement in the discussion". The most negatively assessed items were "Preparation in advance" and "Session Facilitation". "Outcomes of the discussion" was assessed "Very positively" or "Positively" by 37,5% of the respondents.
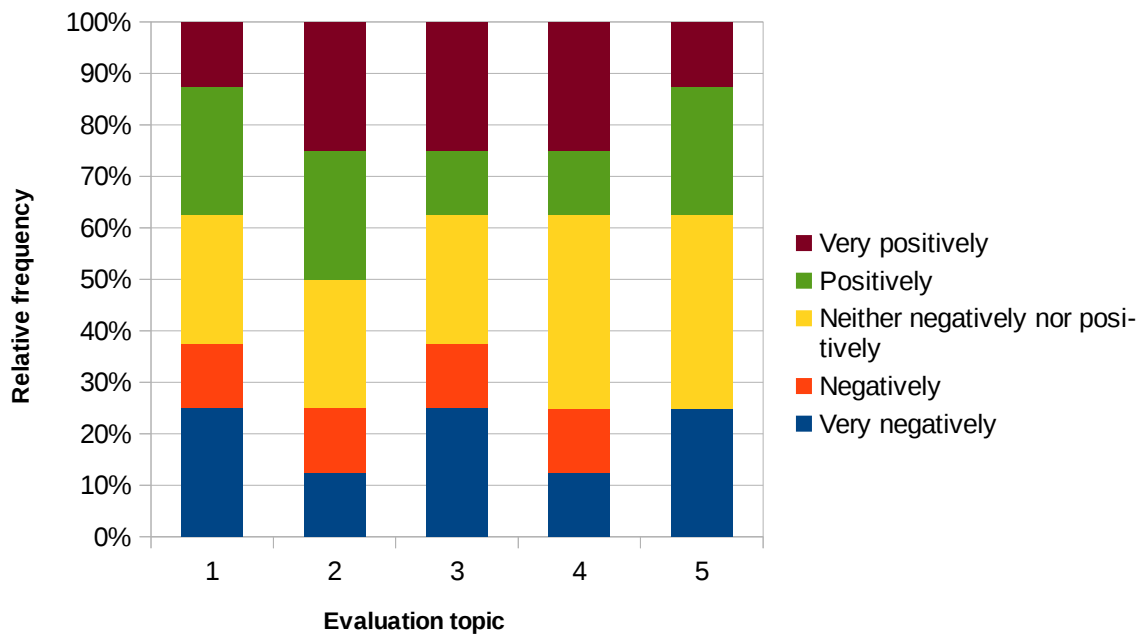
Fig. 7 - "Uses of data, user-science-policy interface and reaching out to the cultural domain" group participants' evaluation of the work session (1 – *Preparation in advance, 2 - Engagement in the discussion, 3 – Session facilitation, 4 – Depth of the discussion, 5 – Outcomes of the discussion*)

5.1.6 . Biodiversity Informatics

Comparing with previous sessions, survey respondents that evaluated the debate on the creation of a Biodiversity Informatics curriculum (N=11) were not, overall very satisfied, with at least 16,7% (N=2) having assessed "Negatively" or "Very negatively" 4 of the 5 items. The one aspect of the session that did not register any "negative" or "very negative" assessments was "Session facilitation" which, nevertheless obtained the highest percentage of "Neither negative nor positive" assessments (44,5%). This aspect, did, however, register the highest levels of "Positive" or "Very positive" ratings - 54,5% (N=6) exactly the same result as that of "Depth of the discussion".
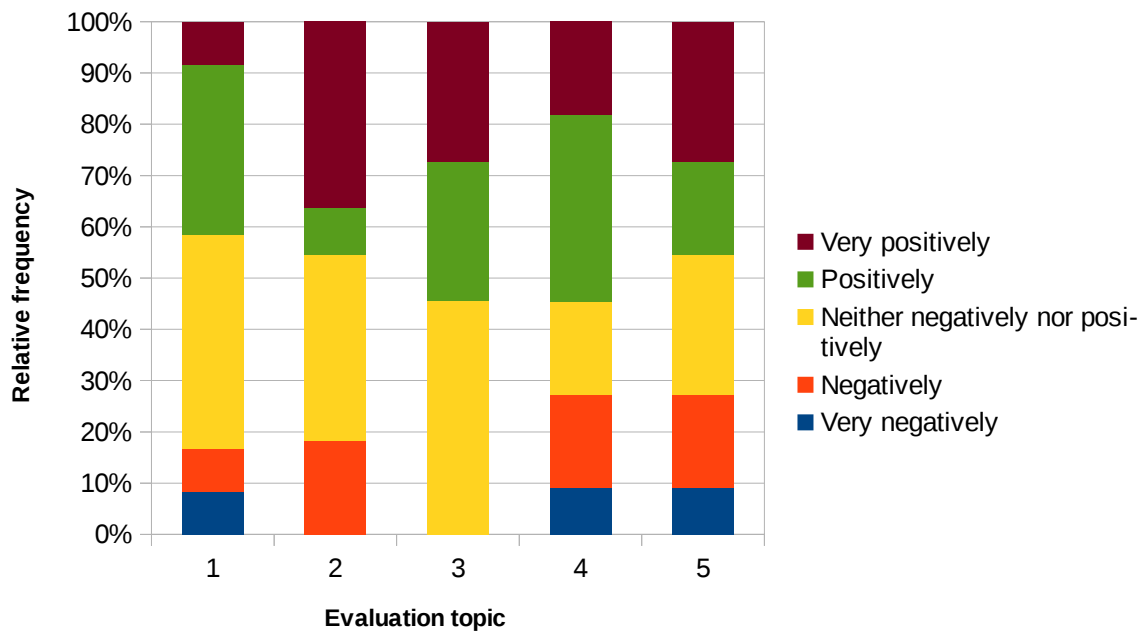
Fig. 8 - Participants evaluation regarding the "Biodiversity Informatics Curriculum" debate session (*1 – Preparation in advance, 2 - Engagement in the discussion, 3 – Session facilitation, 4 – Depth of the discussion, 5 – Outcomes of the discussion*)

5.1.7 . Digitisation and data repatriation

Like in the previous debate session, there were no "Negative" and "Very negative" assessments for all but one of the evaluated aspects. In this session, "Preparation in advance" was the exception which, nevertheless had the lowest percentage of "Positive" or "Very positive" assessments - only 27,2% (N=3 of a total of N=11), the same result as that of "Outcomes of the discussion". On the opposite extreme "Engagement in the discussion" was assessed "Positively" or "Very positively" by 63,7% (N=7) of the respondents.
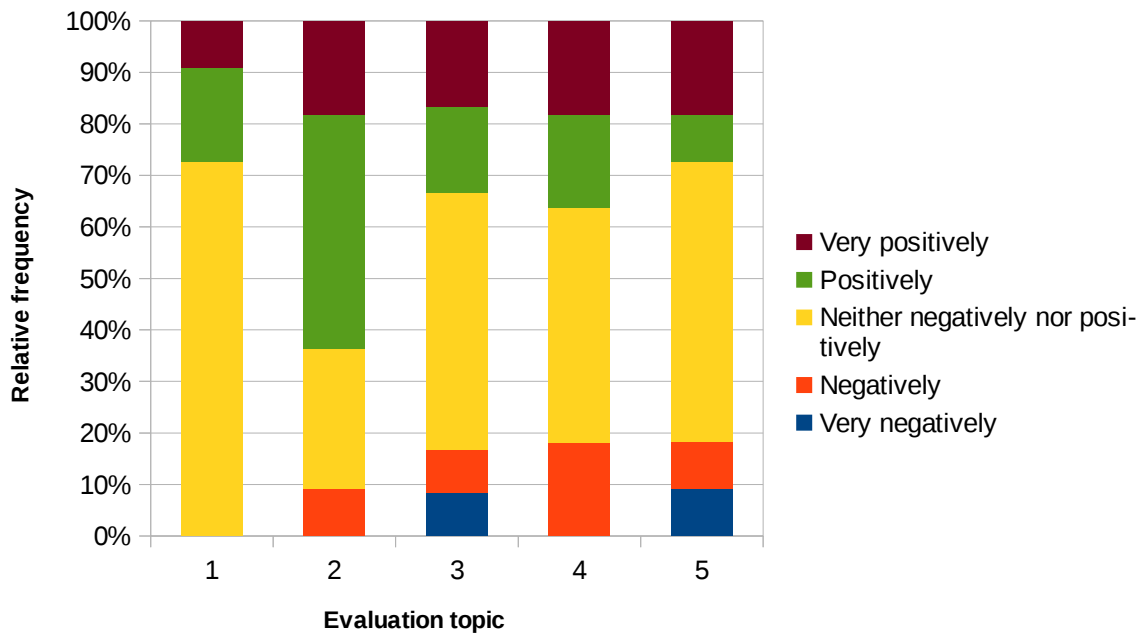
Fig. 9 - Participants evaluation regarding the "Digitisation and data repatriation" debate session (*1 – Preparation in advance, 2 - Engagement in the discussion, 3 – Session facilitation, 4 – Depth of the discussion, 5 – Outcomes of the discussion*)

5.1.8 . Strategic aspects

The assessments of the discussion of the three strategic aspects addressed at the meeting were, overall, very similar and favourable with 46,7-66,7% (N=7-10) of the respondents assessing the debate either "Very positively" or "Positively". However, there were "Negative" assessments regarding the aspects of "European Nodes contribution to the launch of a new fitness-for-use task group on Invasive Alien Species" (20%, N=3) and "New GBIF Members in the European region" (64,3%, N=2).
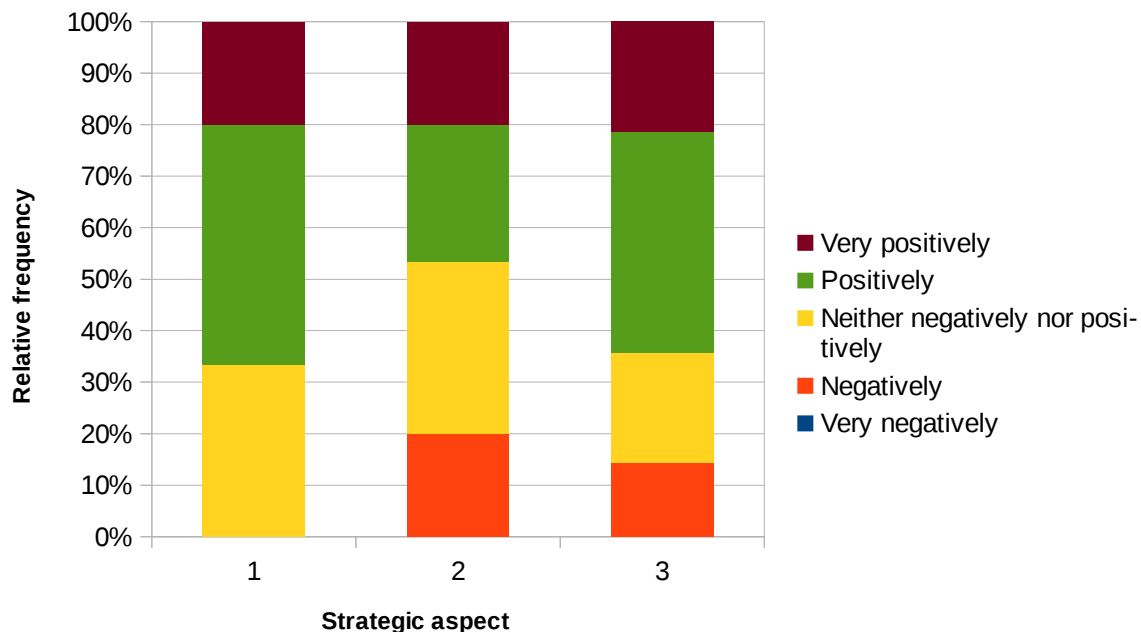
Fig. 10 - Participants evaluation regarding the Atrategic aspects" debate session (*1 – Preparation in advance, 2 - Engagement in the discussion, 3 – Session facilitation, 4 – Depth of the discussion, 5 – Outcomes of the discussion*)

5.1.9 . Final comments

Besides the comments regarding the Venue already mentioned, two of the survey respondents that filled-in the "Comments/Suggestions" field of the Meeting Evaluation Form also pointed:

- "Preliminary work (e.g. agenda) should better be spread between all nodes managers. Assessment of last year results could be improved."

- "Most nodes struggle with a lack of opportunity and resources for progress. Most nodes have not yet reached a stage of 'functional node' - meaning a national biodiversity facility able to support biodiversity data publication and use - including data publishers and data users from outside of the host institute. Most nodes instead experience cuts in the node budget and reduced opportunities to deliver on expected tasks. The progress on the Nodes work-plan is ambitious and in a good direction, but might lack in realism based on the capacity and resource status at the nodes.

  Excellent venue and organization! Even if a meeting in a big city such as Lisbon is very convenient, the spread of participants across many hotels far apart makes for less interaction in the evenings. However, the suggested restaurant was a good idea."

5.1.10 .   Discussion and Conclusions

The Meeting participants that submitted their evaluation of the event were, overall, pleased with the Venue and the practical aspects of the organization, although there were some issues regarding the acoustics of the room (people's voices were muffled) which hampered the discussion. Nonetheless, the goals of the meeting were, according to most of the respondents, either partially or completely accomplished.

As for the discussion of the different subjects, we should make a distinction between breakout group work sessions, among which there were considerable differences, and plenary discussions, which were more even. Among the first, the "Sample-based data publication" and "Nodes portals (and other tools)" group work sessions were the most successful both considering the outcomes of the discussion (most of the participants assessed them "Positively" or "Very positively") and other practical aspects, for which there were no "Negative" or "Very negative" assessments. The same does not hold true for the other two group work sessions: for the "Data quality, validation" group discussion, although there no "Negative/Very negative" assessments, only 25% of the participants assessed the discussion outcomes as "Positive", while the "Uses of data, user-science-policy interface and reaching out to the cultural domain" group, besides assessing "Negatively" and "Very negatively" all aspects from "Preparation in advance" to "Depth of the discussion", also rated poorly the "Outcomes of the discussion", with "Positive"/"Very positive" assessments coming from less than 40% of the respondents.

Regarding plenary discussion sessions, both "Biodiversity informatics curriculum" and "Data Digitisation and Data repatriation" debates had positive and negative assessments regarding different aspects of the discussion and its outcomes, thus showing room for improvement.

Finally, the strategic aspects were also evaluated favourably, with more positive than negative assessments for both the "Admission of New GBIF members in the European Region" and "Definition of the European Nodes Contribution to the launch of a new fitness-for-use task group on Invasive Alien Species" and no negative assessments at all for "Shaping of the European Nodes Work plan 2016-17".

5.2.   Workshop Evaluation Survey Results

A total of 13 participants of the workshop, accounting for 76,5% of the attendees of the event (N=19, excluding 2 workshop organisers plus 2 meeting organisers) submitted their evaluation by completing the online questionnaire.

5.2.1 . General characteristics

92,3% of the respondents said they were either "Very satisfied" or "Satisfied" with the workshop and of the nine general characteristics assessed only two had negative assessments, namely "Venue and equipment" and "Methods/approach", with one participant reporting to be "Dissatisfied" regarding both. "Structure and schedule", "Interaction with peers" and "Practical organization" were the most successful aspects of the workshop, with "Very satisfied" respondents accounting for 38,5%-69,2% of the total, with the remaining reporting "Satisfied".
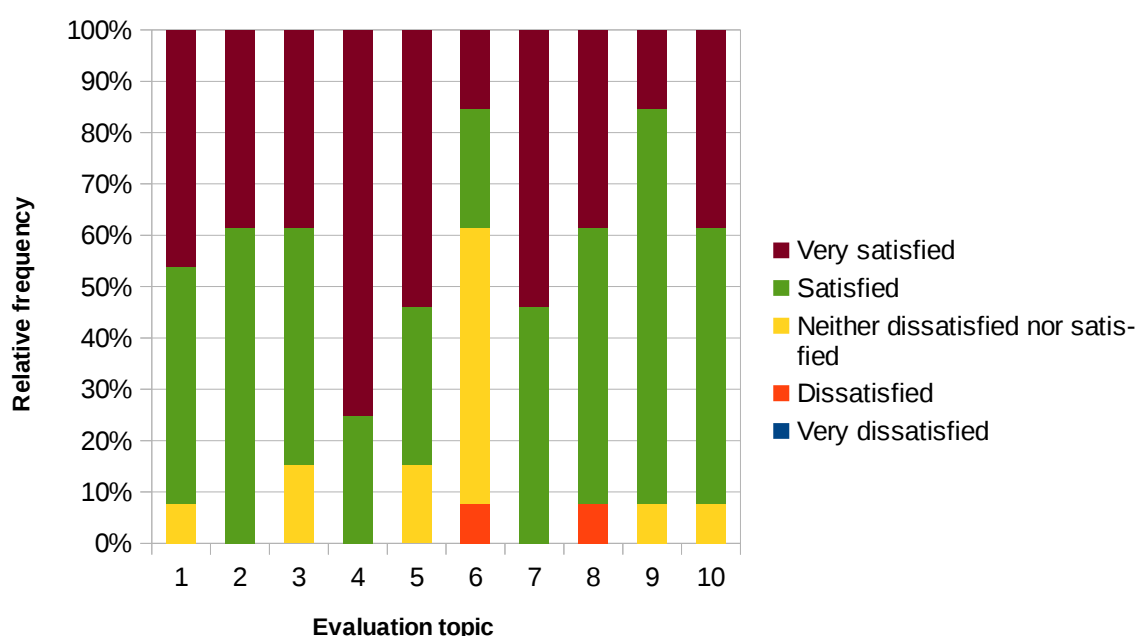

Fig. 11 - Participants' satisfaction levels regarding the Workshop general characteristics(*1 – Content, 2 - Structure and schedule, 3 – Facilitation, 4 - Interaction with peers, 5 - Pre-workshop administration, 6 - Venue and its equipment, 7 - Practical organisation, 8 – Methods/approach, 9 - Results and applicability in work place, 10 - Overall success*)

5.2.2 . Workshop stages

Again, almost all the respondents (N=12) were either "Satisfied" or "Very Satisfied" with the each of the workshop stages: only one participant reported "Neither dissatisfied nor satisfied" regarding "Introduction and introductory presentations", "Group work sessions" and "Wrap up and Conclusions", while "Results presentation and discussion" was satisfactory for 69,2% of the respondents and very satisfactory for the remaining. The stage with more "Very satisfied" assessments was "Group work sessions" (N=8, 61,5%).
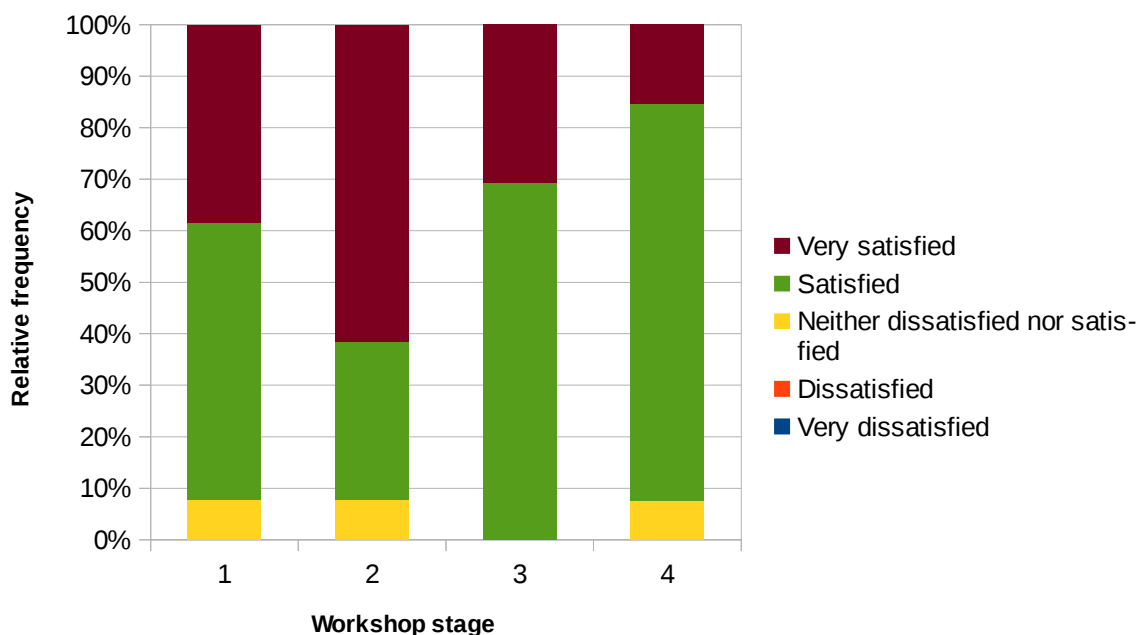
Fig. 12 - Participants' satisfaction levels regarding the Workshop stages (*1 – Introduction and introductory presentations, 2 - Group work sessions, 3 - Results presentation and discussion, 4 - Workshop wrap up and Conclusion*)

5.2.3 . Most useful for daily work

When asked, in a previous question, about the applicability of the workshop results in their workplace, 76,9% of the respondents said they were "Satisfied" and 15,4% "Very satisfied", and here is what they found was most useful for their daily work:

- "Insight into opportunities and challenges of EventCore"
- "Clarification of events hierarchy and use of GUIDs"
- "Improving my knowledge and discussion on Sample Based data and vegetation plots. I will publish the first vegetation data very soon"
- "Improved comprehension on how to deal with different dataset types, resulting from the discussion around all datasets provided by participants. Better understanding of current limitations of DwC schema to support all kinds of sample datasets"
- "Connection and exchange with other GBIF nodes"
- "Practical exercises on DwC mapping with the event extension (choice of the most adequate DwC terms, analysis of the dataset in order to select the best fields to share...)"
- "good introduction to the important steps to prepare and publish sample-event base datasets"
- "Understanding GBIF Nodes work in general, first glance on sample-based data publication"
- "different types of uses case was discussed, which will be really useful for the future"

5.2.4 . Individual datasets

A few weeks before the workshop participants were asked, by e-mail, if they had one or more sample-based species datasets that could be used as real-case examples of the issues that might come up when trying to publish the information through GBIF using the new Darwin Core extension for sample data. The workshop organisers told participants that this could help them and other colleagues make the most out of the workshop. Although only five of the 13 respondents (38,5%) provided datasets, 80% of them (N=4) confirmed it was helpful. Of the other eight participants of the workshop that submitted their evaluation, half said there was "Not enough time" to get a dataset while the other half said that one was "Not available".
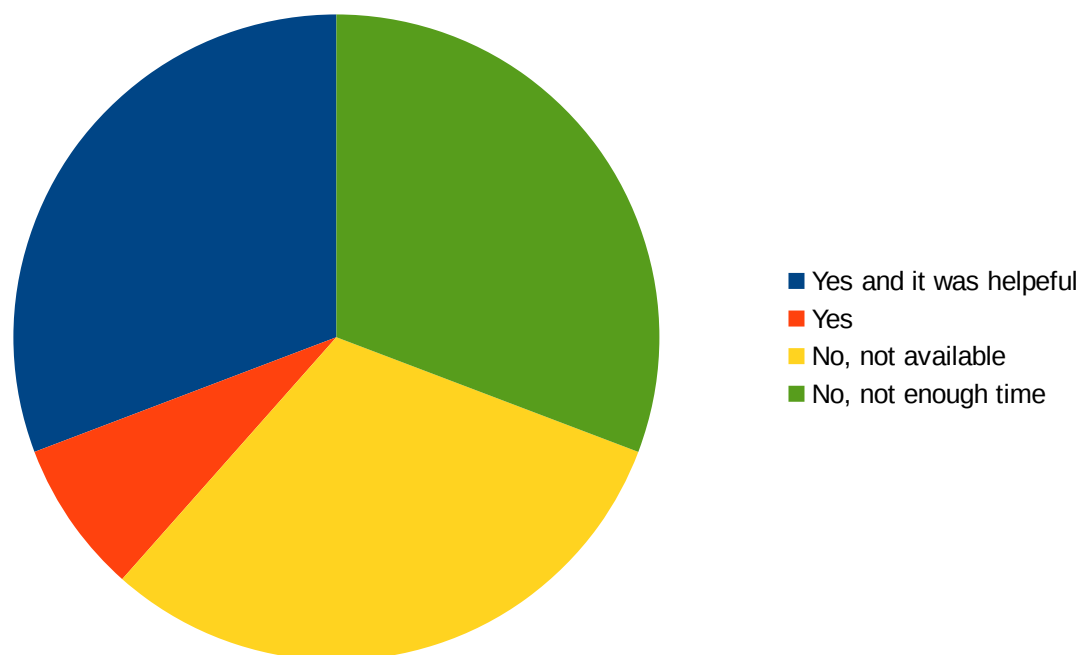

Fig. 13 - Participants' answers to the questions: "Did you provide example sample event datasets for the group work sessions? If so, did you find it helpful? If you didn't, what was the reason?"

5.2.5 . Additions to the workshop and need for Follow-up

Only six respondents answered the question "Is there anything that you would have added to this workshop? Any follow-up after it?" Two of them simply answered "No", while the rest said:
- "Would be good to encourage better documentation of dataset transformation steps : from original files to DwCA"
- "Maybe more time spent with the other groups (apart from the results presentation part), to have a better view of the issues and solutions encountered by the others"
- "Summary of recommendations from workshop, is good product to improve Documentation on how to publish event data"
- "No, it was quite complete and I really hope and think that the follow-up will be interesting for lots of nodes"

5.2.6 . Workshop as a side event of the Nodes Meeting

All 11 respondents that answered the question "Do you recommend continuing the practice of organizing a workshop together with the regional European GBIF Nodes meetings? Lessons learned from this workshop or suggestions for topic for a nodes workshop next year?" think the workshop is an important side event of the Nodes Meeting. Here are the individual answers:
- "Absolutely, for me this was the most useful part of the meeting"
- "Yes, suggested topic: some practical hands-on on GUIDs implementation"
- "Definitely"
- "I do! A set of recommendations spawned from discussions is always a good outcome of a workshop. Stick to that idea."
- "+ The existence of technical workshops between is very important for the nodes, helping them to improve their level and services to their communities
- + Optimization of mission costs
- - For those participants that attend simultaneously to the workshop and to the meeting, the preparation time for both events turns out to be very limited"
- "Yes"
- "Yes. It is an important meeting to keep up in order to give th Nodes coordinators the opportunity to exchange about what is going on and to have the opportunity to discuss topics of international importance and coordinate common priorities at the european level. It is a great opportunity to brainstorm on and initiate collaborative project ideas"
- "Absolutely, it's a great way of working together to solve shared issues among data managers"
- "YES, I think it is the best way to improve the quality of published data, I mean do it right, with direct feedback from the experimented providers"
- "It's a good idea to organize a workshop together with the regional GBIF Nodes meetings"
- "Yes definitely, but only of course if we define a relevant topic for EU nodes"

5.2.7 . Final comments

Three of the respondents made the following comments/suggestions:
- "It was great, unfortunately the sound in the room was horrible, making it reasonable difficult to understand the parlé"
- "It would have been nice to get more information on how sample-based data will be represented on the GBIF data portal"
- "the participants of the workshop bring a lot of energy and I think it was really productive. But we could organise another meeting to follow-up because there is still a lot of question around this topic"

5.2.8 . Discussion and Conclusions

Results indicate that satisfaction levels amongst respondents were high regarding every aspect of the workshop but the Venue (one final comment suggests that the problem was, again, the acoustics of the room). Likewise, participants were pleased with each of the workshop stages and they said that the workshop helped them in different ways after going back to their workplaces. Finally, almost everyone that provided datasets for the group works sessions said they had

benefited from it. Therefore, it is not surprising that respondents were adamant that the workshop is a very important, if not essential, complement to the Nodes Meeting.

6. Event Communication

As part of the organisation of the Meeting, a Communication plan was devised in order to promote the event both inside and outside the GBIF community. In fact, although only GBIF collaborators could attend the meeting, the event was nevertheless considered an opportunity to raise the public profile of GBIF in Portugal and promote its resources amongst potential users, both at the individual level (researchers) and institutional levels (organizations). We used several approaches/tools to reach the different types of audience.

6.1. Meeting website

A conference website was created inside GBIF Portugal's institutional website with the URL: http://www.gbif.pt/EuropeanNodesMeeting. This platform became the main interface used for communication between the organization and the members of the GBIF community in Europe, potential participants (and, to a lesser extent, other audiences). In fact, all the relevant information about the event was published in different sections of the website listed in the Navigation bar (Programme, Venue, Tourist Information, Workshop and Excursion). Besides, the Registration of participants was done through an online form in the Registration section and the Meeting Agenda was discussed online by participants prior the event by means of a GoogleDocs document, accessible through a hyperlink in the Programme section of the website.

The website, in the above address, as well as linked resources (files, GoogleDocs), with be persistently preserved and kept accessible online for future reference.

Fig. 14 - Homepage of the 8th European Nodes Meeting website, accessible at http://www.gbif.pt/EuropeanNodesMeeting.

6.2. Press releases

The Media are an important information vehicle when it comes to the public communication of science. Therefore, two Press Releases (see Annex IV) were prepared and sent out through Instituto Superior de Agronomia's Media Office.

The first Press Release - "**GBIF Portugal organiza 8ª Reunião de Nós Europeus do GBIF em Lisboa**" ("GBIF Portugal host 8th European GBIF Nodes Meeting in Lisbon") - was e-mailed to portuguese Media, two weeks before the event, on April 8th, while the second one - **"Informática da Biodiversidade: Nós Europeus do GBIF discutem em Lisboa criação de currículo"** ("Biodiversity Informatics: European GBIF Nodes discuss creating a curriculum in Lisbon") - was e-mailed during the Meeting, on April 21st.

6.3. E-mailing

In order to reach Natural Sciences researchers and other professionals that may use GBIF's biodiversity data, we compiled a list of Natural Sciences Academic/Research Institutions as well as Natural History Museums and Botanical Gardens, Scientific associations and Environmental/Biodiversity NGOs in Portugal and, on April 13th, an e-mail was sent out to 68 institutions/organizations (see Annex V) requesting the dissemination, in their communities, of the Press release announcing the hosting of the 8th European GBIF Nodes Meeting by GBIF Portugal. A Scientific society and an NGO posted about the event on their Facebook timeline and Instituto Superior de Agronomia and Fundação para a Ciência e Tecnologia, which had been sent personalised requests a few weeks before, also acted upon them by posting the event information on their websites(see Annex VI).

6.4. Social Media

GBIF Portugal is present in two Social Media platforms where the Meeting was advertised before and during the event.

6.4.1 . Twitter

A total of 9 tweets were posted in GBIF Portugal's timeline at https://twitter.com/gbifportugal.

Fig. 15 - 1st tweet about the Meeting

Fig. 16 - 2nd and 3rd tweets about the Meeting

Fig. 17- 4th tweet about the Meeting

Fig. 18 - 5th tweet about the Meeting

Fig. 19 - 6th tweet about the Meeting

Fig. 20 - 7th tweet about the Meeting

Fig. 21 - 8th tweet about the Meeting

Fig. 22 - 9th tweet about the Meeting

6.4.2 . Facebook

One post appeared in GBIF Portugal's timeline at **https://www.facebook.com/N%C3%B3-Portugu%C3%AAs-do-GBIF-294265787393100/**.

Fig. 23 - Facebook post about the Meeting

Annex I.
8th European Nodes Meeting - Agenda

8th European Nodes Meeting

19-21 April 2016, Lisbon

<u>Day 1: Tuesday 19 April (2-6PM)</u>

(Plenary session)

14:00 - 14:20

1. **Opening session by the President of FCT (t.b.c.), the President of ISA, Head of Delegation from Portugal, Director of RI PORBIOTA and GBIF Portugal**

14:20 - 14:50

2. **Introduction**

> Tour de table
>
> Preparatory Survey (Anne-Sophie Archambeau)
>
> [Presentation - Survey.pptx](#)
>
> Setting up the scene (Anne-Sophie Archambeau)
>
> Nodes and regional work plans and activities (Hanna Koivula)
>
> [Presentation - NC_and_Regional_WP.pptx](#)

Questions

14:50 - 16:00

3. **Update from the Secretariat (Kyle Braak and Alberto González)**

> 1. Strategic Plan 2017-2021
>
>   [Presentation - GBIF Strategic Plan 2017-2021 - small.pptx](#)
>
> 2. 2016 update
>
>   1. Content mobilization and quality
>
>   - Sample-based data mobilization
>   - Task groups
>
>   2. Informatics
>
>   - Licensing
>   - Data publisher and user agreements

- National portal developments (ALA)

3. Communications and capacity enhancement

- New documentation
- Capacity Enhancement Support Programme (CESP)
- Capacity self-assessment
- Country reports and regional pages
- Certification of skills (badges)
- BID and BIFA programmes

Presentation - GBIF Update Slides For Regions - small.pptx

Discussion

16:00 - 16:30 **Coffee break**

16:30 - 18:00

4. Collaborations & reporting and new opportunities (max 10'- just an update)

1. EUBON (max 5'- just an update) (Anders Telenius)

2. LifeWatch (max 5'- just an update) (Anders Telenius + Rui Figueira)

Presentation - EU%20BON%20LifeWatch%20ENM%202016.pptx

3. COST Action (first try turn down but next step?) (Anne-Sophie Archambeau)

Presentation - COST-Presentation.pptx

4. EU Directives and Funding opportunities (André Heughebaert)

5. EASIN (Anders Telenius)

6. CoopBioPlat (Cristina Villaverde)

Presentation - CoopBioPlat-CV.pptx

7. Report Mentoring Encounter Bay (on ALA documentation) (André Heughebaert)

8. CESP GBIF_PALOP (Rui Figueira)

Presentation - GBIFPALOP_RuiFigueira.pdf

9. Citizen Science providers in Europe and data availability (Nils Valland)

Presentation - European_CitizenScience_compiledNBIC_20160418_0.pptx

Speadsheet - CitizenScience_SpeciesOccurence_DataAvailabilityEurope_20160219_NilsValland

[%40NBIC_0.xlsx](#)

   10. AnnoSys - Annotation system for biodiversity data portals (Jörg Holetschek)

      [Presentation - AnnoSysIntegrationInGBIFdataPortal.pptx](#)

Wrap up discussion

Day 2: Wednesday 20 april (9AM-6PM)

In order to follow up with our [EU joint effort work plan 2015-2016](#) and the [global nodes meeting](#), here are the different topics that we will discuss during the regional meeting. The links goes to the community web site pages of the existing groups.

   5. **Best practise exchanges organized in thematic groups**

9:00 - 9:15 **Introduction to working groups**

9:15 - 10:55 **Parallel session 1: 2 thematic groups**

A - Nodes portals (and other tools) (ALA) (AS Archambeau)

      [http://community.gbif.org/pg/pages/view/49268/work-plan-group-green-sharing-and-jointly-developing-nodes-portal-solutions](#)

      NOTES from the session:

      [https://docs.google.com/document/d/1UuvXbfrU73MurfVntWgqihDbMi4jcYWYElcAxts-ZBg/edit?usp=sharing](#)

B - Sample based data (Dag, Kyle)

      [http://community.gbif.org/pg/groups/47949/samplebased-data-publishing-interest-group/](#)

      The pre-meeting workshop on sample-based data: [https://goo.gl/oqg2OR](#)

10:55 - 11:20 **Coffee break**

11:20 - 13:00 **Parallel session 2 : 2 thematic groups**

C - Uses of data, users-science-policy interface (Dimi and Liam), reaching out to the cultural domain

(EUROPEANA, Jörg Holetschek)

http://community.gbif.org/pg/pages/view/49063/

http://community.gbif.org/pg/pages/view/49334/work-plan-group-red-improving-relevance-of-gbif-for-sciencepolicy-initiatives

Example data use template:

https://docs.google.com/document/d/1AZ11Axu747H7e386R6KaIC4FjWoSbdjn3V4kSzAp3oE/edit?usp=sharing

NOTES from the session:

https://drive.google.com/open?id=15D5qwWQcnY0oirpx8kIS11iTqCuxCCuq-CvI1qEvg0U

D - Data quality, validation (and fitness for use) (Christian and Dag)

http://community.gbif.org/pg/pages/view/49262/work-plan-group-blue-improving-the-quality-of-gbif-mediated-data

Notes from the 11th EU Nodes meeting data quality WP group:

NOTES from the session:

https://docs.google.com/document/d/197Btzw4CXgQl6d6Bjcsb63wNbXQ2yoTO_HHvKZfRjmo/edit

GBIF Europe GitHub: https://github.com/GBIF-Europe

13:00 - 14:00 - Lunch

14:15 - 15:00 **Plenary session**

    1. Presentation about Europeana (Jörg)

      Presentation - OpenUp.pptx

    2. Summary about the sample data workshop (Dag)

    3. Wrap up of thematics groups (15' each group)

      1. Nodes Portal (Anne Sophie)

        Notes from the group meeting

      2. Sample data (Dag, in combination with the workshop report)

[Workshop report](#)

3. Users and science-policy interface (Dimitri)

[Notes from the group meeting](#)

4. Data quality (Dag)

[Notes from the group meeting](#)

15:00 - 16:00 **Discussion around the next Action plan (COST) including aspects on:**

1. Biodiversity informatics curriculum (Hanna)

http://community.gbif.org/pg/pages/view/48692/breakout-group-towards-a-curriculum-for-biodiversity-informatics
[Notes from the meeting](#).

2. Digitization and repatriation (Anders)

http://community.gbif.org/pg/pages/view/48970/breakout-group-common-approaches-to-digitization-and-repatriation

[Presentation - GNM 13 B-O Yellow Digitization and Repatriation.pptx](#)

3. Extending checklists, to include National Red list, black list, crop wild relative list, evidence, new data attributes needed (Wouter Addink?)

4. Data mobilization targets, known missing datasets, who will approach data owners…? Data hosting services?

16:00 - 16:30 **Coffee break**

16:30 - 18:00 **Following of the discussion and Wrap up**

19:00 Hosted dinner and visit to Jardim Botânico da Ajuda

Day 3: Thursday 21 april (9AM-6PM)

Plenary session

9:00 - 10:30 **Strategic aspects :**

Regional issues and correlation with strategic plans (Hanna)

1. European Nodes Workplan 2016-2017

Defining goals, Priorities and Actions for the coming 12 months

1. Develop processes and strategies to advance GBIF-related activities within the region

2. Strengthen sharing of capacities and resources within the region

3. Promote engagement and expansion of the GBIF network

Looking at EU nodes Work Plan of last year, what works and what doesn't work?

http://community.gbif.org/pg/file/read/48045/european-nodes-work-plan-20152016

--

European Nodes WP actions: TABLE in GoogleDocs

10:30 - 11:00 - **Coffee break**

11:00 - 13:00 **Following of the discussion**

13:00 - 14:00 - **Lunch**

14:00

Presentations of GBIF activities in Russia by Maxim Shashkov

Presentation - Shashkov_Ivanova_MSh_NI.pdf

Presentation of Zoological Institute RAS by Roman Khalikov

Presentation - Khalikov_GBIF.pptx

Following of the discussion

Possible launch of a new fitness-for-use task group on IAS. How can we contribute?

Interaction with HoD

New members Russian GBIF and other possible new members related to LifeWatch

Defining the actions

2. **Wrap up and Conclusion**

Maybe of interest links:

http://www.openaccess.nl/en/events/opening-the-book-on-open-access

CAPACITY SELF ASSESSMENT FOR GBIF PARTICIPANTS:

http://www.gbif.org/resource/82277

Annex II.
Workshop on Sample-based Data Publication – Agenda and preparatory information

Workshop on sample-based data publication

18-19 April

Organized by Dag Endresen (GBIF Norway), Kyle Braak (GBIF Secretariat), Christian Svindseth (GBIF Norway) and Anders G. Finstad (GBIF Science Committee, NTNU University Museum in Trondheim)

Same venue as the EU Nodes meeting at Instituto Superior de Agronomia (ISA).

http://www.gbif.pt/EuropeanNodesMeeting
http://www.gbif.pt/EuropeanNodesMeeting/workshop

This page has a shorter URL: https://goo.gl/oqg2OR

Shared Google Drive folder for the workshop: https://goo.gl/IY1YUx

Our primary aim for the Nodes workshop is for the participants to form groups and discuss specific aspects and issues with the publishing of sample-based, event-core data and datasets - rather than IPT training of the same type as in Madagascar. Our aim is for the sub-groups to identify issues and to suggest possible improvements for this new data type. Most of the groups might choose to work directly with the IPT to explore issues and solutions, other sub-groups might choose other approaches. We will provide a demo IPT for the workshop.

We suggest here some topics for the groups and invite all nodes and participants to add additional topics and issues related to publication of sample-based data. Workshop participants should already in advance start to select/vote for topics of interest.

Preparations before the workshop:

Please contribute to the list of issues and topics on sample-based data for the sub-groups. Start to evaluate which sub-group you wish to join. Identify datasets from your own node to illustrate the sub-group topics with specific issues to be discussed and if possible solved during the workshop. Please identify and provide several example data sets for multiple topics.

Agenda

Monday 18th April

09:00 Welcome and introduction to the workshop (Dag Endresen & Kyle Braak, 15 minutes)

09:30 Introduction to what is sampled-based data (Anders Finstad / Dag / Kyle)

10:00 Experience with sample-based data at GBIF Norway including processing (R) of datafiles for publication using the IPT (~~Christian Svindseth~~ / Anders Finstad)

10:15 [Presentation](#) on EASIN (Ana Cardoso)

10:30 Coffee break

11:00 Progress on handling sampled-based data at the GBIF secretariat (Kyle Braak)

11:30 Presentation of sample-based data in the Norwegian dataset portal (~~Christian Svindseth~~)

12:00 Brief presentation of each sub-group topic and registration of participants to sub-group

13:00 Lunch break

14:00 Split in sub-groups (to work with the dataset examples, and sampling data topics)

16:00 Coffee break

17:00 End of day 1

19:00 [Ice-breaker](#)

Tuesday 19th April

09:00 Each sub-group wrap-up and to prepare for plenary presentation of results

10:00 Sub-group presentation and discussion in plenary

10:30 Coffee break

12:00 Summary of workshop results (Dag, Anders and Kyle)

13:00 Lunch break

14:00 The main EU Nodes meeting start ([agenda](#))

Some general questions

What is sample-based data?
How do we recognize if a dataset is sample-based?
Can the same dataset be represented using different core types?
How to present and visualize sample-based data in the GBIF portal?
How do we best model sample-based data for reuse in new contexts?
How to design sample-based data APIs and web services.

Some outcomes to report on:

- Were you able to find out enough information (metadata) about the dataset to understand it?

Could you fill in the missing columns in the [Overview - Example Datasets spreadsheet](#)?

- Which issues (topics below) did you encounter while working on the example dataset?
  - What format (event vs. occurrence) to use, also which extensions
  - How to handle presence / absence data
  - How to record the sampling protocol, abundance, other DwC fields
  - How to transform the raw data into DwC tabular form
- For each issue you encountered, did you find a solution / recommendation? Please descrIbe.
- Did you successfully publish the example dataset into the sample event format?
- What documentation did you use to learn more about how to publish sample event data? E.g. what DwC terms to populate, how to populate DwC terms, list of extensions available for use, etc.
- Was the documentation sufficient, or was there something missing? E.g do you need these documentation in your local language, or is English sufficient?

Suggested topics / issues for sub-groups

(Please add and suggest other topics!)

1. *Occurrence* data versus *sample-based* data
When is the different data type core appropriate? Can the same data be represented in both cores? Are there benefits or issues to take notice of here? Data previously published as occurrence-core data could be published again (improved) as event-core data?

2. Events and occurrences published with separate datasets
Splitting the description of sampling events and sampled occurrences into separate datasets when publishing in GBIF. Limitations of star-schema and extensions with [measurementOrFact](#). How OBIS is working around this limitation with their "[Extended Measurement or Facts Extension](#)". Material collected during the very same sampling event can be stored by different institutions who publish their own separate datasets in GBIF. Split projectID, instituteID, collectionID etc into a "*GBRDS*"-like system. Or it can be stored in the same dataset (see Example dataset Lepidoptera Collection by Hannu Saarenmaa below).

3. Visualization of sample-based data
Presentation and visualization of sample-based data in the GBIF web portal (Basic ability to search/filter datasets by type). Add Events as entry point in the GBIF portal? Add other new data types as entry points?
Example demo: http://data.gbif.no/datasets/events/events

4. Multiple occurrence evidence types for the same sample
Issues with more than one source of "evidence" for the same occurrence. One plant or animal sampled

for herbarium, tissue for DNA, etc. Human observation, specimen and multimedia evidence for the same occurrence.

5. Absence data

Explicit documentation of absence data (i.e. species not observed on a particular sampling event) for surveys with a large number of target species (e.g. all vascular plant species present in sample plot), sampling effort, sampling method. Display and visualisation of absence information - not show absence as occurrence points. Taxon coverage attribute added for the event core?

6. Abundance, population size

How to document abundance type data. We need to describe the type of sampling gear, the size of the sampling gear, number of replicates, the organism amount (and in addition often the biomass) in each sample. The importance of samplingProtocol, sampleSizeValue + sampleSizeUnit, samplingEffort and organismQuantity + organismQuantityType individualCount.

7. New data types

Add determination keys, etc...

Measurement and facts, traits

Multimedia, http://www.gbif.org/occurrence/search?MEDIA_TYPE=*

8. Verbatim description of data

Example: "*Number* of trout sampled using *such* type of …"

9. Transforming crosstab source data formats

Transforming crosstab format to Darwin Core archive format in lists, R-script, R-package, software tool hackathon, Ruby, Python, ...

10. New data input formats (for the IPT)

To handle sample event data, should new data input formats be supported by the IPT, e.g. abundance matrix.

11. Sampling methodology, sampling protocol

For a large projects using a variety of biotic and abiotic sampling protocols, should all events be captured, only the biotic ones that allow us to monitor the changes to the biodiversity over time?

12. Grouping events (using parentEventID)

The parentEventID needs to relate to an event record in the same dataset with that eventID, otherwise it's a broken relation. An exception could probably be made on when the parentEventID is a globally unique identifier that resolves to the parent event (e.g. DOI, HTTP URI, etc).

A child event should (ideally) have the same date and location as the parent event. The classic use is a sub-sampling of a larger parent plot). On the other hand if an event is part of a series of events at the

same location, but different date, then it's better to group them by a common locationID instead of using parentEventID.

13. Obfuscating (fuzzifying) sensitive data (e.g. threatened species location)
Care has to be taken to obfuscate the species location, and the location of the sample event from which it was derived! In Dutch Vegetation Database, threatened species are obfuscated to 5x5 km grids. In DwC, dwc:dataGeneralizations, dwc:georeferenceRemarks can be used to explain how the data was obfuscated for example.

14. Replacing the existing dataset (occurrence) with richer sample event formatted version
Does it have to become a new dataset? Merged into updated version? Assigned a new DOI? etc

Links

Darwin Core: http://rs.tdwg.org/dwc/terms/
Occurrence core: http://rs.gbif.org/core/dwc_occurrence_2015-07-02.xml
Event core: http://rs.gbif.org/core/dwc_event_2015_05_29.xml

Sample-based data: http://www.gbif.org/newsroom/news/sample-based-data

GBIF NODES data publishing training event, GB22 Madagascar, October 2015
http://community.gbif.org/pg/pages/view/47903/schedule-for-the-gb22-training-event

Data publishing workshop in Trondheim, Norway, October 2015: http://www.gbif.no/events/2015/data-publishing-workshop-october-2015.html

Presentation by Éamonn Ó Tuama, April 2014: Data standards: publishing sample-based data using the GBIF Integrated Publishing Toolkit [Link]

Presentation by Éamonn Ó Tuama, March 2015: Darwin Core for Sample Data [Link]

References

Dengler J, Jansen F, Glockler F, Peet RK, De Caceres M, Chytry M, Ewald J, Oldeland J, Lopez-

Gonzalez G, Finckh M, Mucina L, Rodwell JS, Schaminee JHJ, and Spencer N (2011). The Global Index of Vegetation-Plot databases (GIVD): a new resource for vegetation science. *J. Veg. Sci.* 22: 582-597. doi:10.1111/j.1654-1103.2011.01265.x

GBIF and EUBON (2015). Publishing sample data using the GBIF IPT. Latest version 23 March 2015. Available at http://www.gbif.org/sites/default/files/gbif_IPT-sample-data-primer_en.pdf

Hennekens SM, and Schaminee JHJ (2001). Turboveg, a comprehensive database management system for vegetation data. *J. Veg. Sci.* 12: 589-591. doi:10.2307/3237010

Hoffmann A, Penner J, Vohland K, Cramer W, Doubleday R, Henle K, Koljalg U, Kuhn I, Kunin W, Negro J, Penev L, Rodriguez C, Saarenmaa H, Schmeller D, Stoev P, Sutherland W, O Tuama E, Wetzel F, and Hauser C (2014). The need for an integrated biodiversity policy support process – Building the European contribution to a global Biodiversity Observation Network (EU BON). *Nature Conservation* 6: 49-65. doi:10.3897/natureconservation.6.6498

Morris RA, Barve V, Carausu M, Chavan V, Cuadra J, Freeland C, Hagedorn G, Leary P, Mozzherin D, Olson A, Riccardi G, Teage I, and Whitbread G (2013) Discovery and publishing of primary biodiversity data associated with multimedia resources: The Audubon Core strategies and approaches. *Biodiversity Informatics* 8(2):185-197. doi:10.17161/bi.v8i2.4117

Robertson T, Döring M, Guralnick R, Bloom D, Wieczorek J, Braak K, Otegui J, Russell L, Desmet P (2014) The GBIF Integrated Publishing Toolkit: Facilitating the Efficient Publishing of Biodiversity Data on the Internet. *PLoS ONE* 9(8): e102623. doi:10.1371/journal.pone.0102623

Wieczorek J, Banki O, Blum S, Deck J, Doring M, Droge G, Endresen DT, Goldstein P, Leary P, Krishtalka L, O Tuama E, Robbins RJ, Robertson T, and Yilmaz P (2014). Meeting Report: GBIF hackathon-workshop on Darwin Core and sample data (22-24 May 2013). *Standards in Genomic Sciences* 9: 585-598. doi:10.4056/sigs.4898640

Wieczorek J, Bloom D, Guralnick R, Blum S, Döring M, Giovanni R, Robertson T, and Vieglais D (2012) Darwin Core: An Evolving Community-Developed Biodiversity Data Standard. *PLoS ONE* 7(1): e29715. doi:10.1371/journal.pone.0029715

Wiser SK, Spencer N, De Caceres M, Kleikamp M, Boyle B, and Peet RK (2011). Veg-X – an exchange standard for plot-based vegetation data. *J. Veg. Sci.* 22: 598-609. doi:10.1111/j.1654-1103.2010.01245.x

Example sample event datasets

Search for sample-based dataset published in GBIF:

http://www.gbif.org/dataset/search?&type=SAMPLING_EVENT

Please share the original data files using the shared Google Drive folder for this workshop: https://goo.gl/IY1YUx

Sampled based fresh-water fish and invertebrates from the NTNU University Museum

This dataset provide example for the need to look at how to describe sampling methodology

> NTNU University Museum (2015). Lepidurus arcticus survey Northeast Greenland 2013. Sampling event dataset available from GBIF at: doi:10.15468/ancuku (16 occurrences)

Sampled based vegetation data from the NTNU University Museum

> NTNU University Museum (2015). Lowland lakeshore vegetation in Nord-Trøndelag, Norway. Sampling event dataset available from GBIF at: doi:10.15468/kvxnre (782 occurrences)

> NTNU University Museum (2015). Vegetation data from sheep grazing experiment at alpine site in Hol, Norway. Sampling event dataset available from GBIF at: doi:10.15468/dwxqai (15 973 occurrences)

> NTNU University Museum (2015). Vegetation data from long term sheep grazing experiment in Setesdal, Norway. Sampling event dataset available from GBIF at: doi:10.15468/qegaql (11 961 occurrences)

Maritime collections of benthic invertebrates in Oslofjord and Skagerrak coastal areas, 1950-1955

Sampling locations published in one dataset and occurrences published across other datasets linked to the sampling location using the eventID and locationID.

> UiO Natural History Museum (2015). Dannevig collections. Dataset available from GBIF at: http://doi.org/10.15468/hwvr0m (280 collecting events)

> UiO Natural History Museum (2015). Drøbak collections. Dataset available from GBIF at: http://doi.org/10.15468/mg7l2t (637 collecting events)

UiO Natural History Museum (2015). BIOSKAG collections. Dataset available from GBIF at: http://doi.org/10.15468/mpifue (178 collecting events)

UiO Natural History Museum (2015). Dannevig- and Drøbak collections of Polychaeta. Dataset available from GBIF at: http://doi.org/10.15468/y6cctp (data publication in progress)

UiO Natural History Museum (2015). POLYSKAG collections. Dataset available from GBIF at: http://doi.org/10.15468/zfcaa5 (50 sampling-event occurrences)

Oug E, Christiansen ME, Dobbe K, Rønning A-H, Bakken T, and Kongsrud JA (2015). Mapping of marine benthic invertebrates in the Oslofjord and the Skagerrak: sampling data of museum collections from 1950-1955 and from recent investigations. *Fauna Norvegica* 35: 35-45. DOI:10.5324/fn.v35i0.1944

UiO Natural History Museum (2015). Dannevig collections sampling stations 1950-1953 [field-books]. Document identifier: URN:NBN:no-52235. Available from the UiO DUO Research Archived at: http://hdl.handle.net/10852/48310

Species plots from the Norwegian Vegetation Mapping Programme, NIBIO

Monitoring dataset published as occurrence core. What advantages (limitations?) could publishing this dataset as event core provide?

Norwegian Institute of Bioeconomy Research (NIBIO) (2015). Species plots from the Norwegian Vegetation Mapping Programme, 2015-11-04. Occurrence dataset available from GBIF at: http://doi.org/10.15468/na7jbv

Bryn A, Kristoffersen H-P, Angeloff M, Nystuen I, Aune-Lundberg L, Endresen D, Svindseth C, and Rekdal Y (2015). Location of plant species in Norway gathered as a part of a survey vegetation mapping programme. *Data in Brief* 5: 589-594. DOI:10.1016/j.dib.2015.10.014

Insects from light trap (1992–2009), rooftop Zoological Museum, Copenhagen

Systematic 18-year continuous monitoring of insects, example of long-term time-series published in GBIF as sample-based data. http://www.gbif.org/newsroom/news/18-years-of-moth-data

Natural History Museum of Denmark: Insects from light trap (1992–2009), rooftop Zoological Museum, Copenhagen. Sampling event dataset available from GBIF at: http://doi.org/10.15468/xabmiz (37 477 occurrences)

Dataset provided by publisher:
Karlsholt O (2015): Insects from light trap (1992–2009), rooftop Zoological Museum,

Copenhagen. v1.3. Natural History Museum of Denmark. Dataset/Sampling-event. http://danbif.au.dk/ipt/resource?r=rooftop&v=1.3

Thomsen, P. F., Jørgensen, P. S., Bruun, H. H., Pedersen, J., Riis-Nielsen, T., Jonko, K., Słowińska, I., Rahbek, C. and Karsholt, O. (2016), Resource specialists lead local insect community turnover associated with temperature – analysis of an 18-year full-seasonal record of moths and beetles. *J Anim Ecol* 85: 251–261. doi:10.1111/1365-2656.12452

Example sample event datasets submitted by workshop participants:

Please feel free to add links and descriptions of your own example datasets below.

Jörg Holetschek (GBIF-Germany)

1. Managing open habitats for species conservation: the role of wild ungulate grazing, small-scale disturbances, and scale - Tanacetum vulgare dominated forb vegetation. http://doi.org/10.15468/wxtdrg
2. Managing open habitats for species conservation: the role of wild ungulate grazing, small-scale disturbances, and scale - Pinus sylvestris pioneer forest. http://doi.org/10.15468/baymta
3. Managing open habitats for species conservation: the role of wild ungulate grazing, small-scale disturbances, and scale - Corynephorus canescens dominated grassland. http://doi.org/10.15468/85hs5a

   Original data has been provided: https://drive.google.com/open?id=0B5txXNTwbgH8ZXEtdlJTRlRXV1U

   Sampling methodology: These three datasets represent three successional stages in the same project: (1) Corynephorus canescens-dominated grassland; (2) ruderal tall forb vegetation dominated by Tanacetum vulgare; and (3) Pinus sylvestris-pioneer forest. The study was conducted over 3 yr. In each successional stage, six paired 4 m2-monitoring plots of permanently grazed versus ungrazed plots were arranged in three random blocks. Removal of grazing was introduced de novo for the study. In each plot, percentage cover of each plant and lichen species and total cover of woody plants was recorded.

   The sampling method for all three datasets is "Londo, G. 1984. The decimal scale for releves of permanent quadrats." This sampling method is described in:

   Knapp, R. (ed.) Sampling methods and taxon analysis in vegetation science. pp. 45–48. Junk

Publishers, The Hague, NL. Available for download in Google Drive here: https://drive.google.com/open?id=0B5txXNTwbgH8VS1DMUR6QTYtX1U

A description of the project can be found here: http://onlinelibrary.wiley.com/doi/10.1111/j.1654-109X.2010.01119.x/pdf

Pre-workshop notes: Good case study transforming an existing (very basic) occurrence dataset into a richer sample event dataset. Good case study showing how difficult it is to transform the original raw data into tabular DwC tables. Begs the question of whether it's better to treat these as 3 separate datasets, or whether it's cleaner to include all 3 in a single dataset.

Workshop summary:
- *They (Tania and Jorg) successfully managed to create a DwC-A, with event core, and occurrence extension.*
- *They didn't use the IPT, but created the CSV files and meta.xml manually.*
- *They used the Primer, DwC terms, and rs.gbif.org as resources, but discovered the Primer was out-of-date.*
- *They were able to represent a hierarchy between sampling large plots and smaller sub-plots using parentEventID.*
- *They ignored absence data, but this was just to work around GBIF's inability to handle absence data properly.*
- *They reported confusion about the Releve being a special type of event*
- *They reported confusion about what to enter for organismQuantity for the Londo Scale - currently missing from the GBIF Quantity Type vocabulary.*

Post-workshop follow up:
- *Kyle has asked Jorg if he'd like to replace the version of the dataset on GBIF with the new sample event formatted DwC-A produced at the workshop.*
- *To start with, Kyle encouraged Jorg to validate the DwC-A using the IPT.*
- *[DOCUMENTATION] They recommended having better guidelines for how to enter the sampling protocol (e.g. in the updated Primer)*
- *[STANDARDS] They recommended creating a vocabulary for sampling protocol*
- *[STANDARDS] They recommended using the MoF extension for allowing users to add user-defined Releve layers versus the current Releve extension. Certainly it seems overwhelming to try and make a custom extension for every predominantly used sampling protocol. (+1 from Dimitry B)*
- *[STANDARDS] They recommended new term "treatment" be added to DwC, to be able to enter a treatment intentionally applied during the sampling event (e.g. grazing to a vegetation plot). This is different from dwc:preparations, which applies to the specimen.*
- *[STANDARDS] They recommended adding the Londo Scale to the GBIF Quantity Type vocabulary. (+1 from Dimitry B)*
- *[VISUALISATION] They expressed a desire for a browser capable of visualising the*

event hierarchies

Andre Heughebaert (Belgian Biodiversity Platform)

1. ManscapeData: Integrated management tools for water bodies in agricultural landscapes. Original data provided, but must remain private. For more information, please refer to the final report: https://drive.google.com/open?id=0B5txXNTwbgH8QTZ3amY3YkVSNms and a scientific article derived from this dataset: https://drive.google.com/open?id=0B5txXNTwbgH8TGswZXIwcnM4bzQ

   Sampling methodology: During the summer of 2003 (end of July until beginning of September), we simultaneously surveyed a total of 99 ponds for a number of physical and chemical water parameters and communities of 12 groups of organisms: bacterioplankton, phytoplankton, benthic diatoms, rotifers, cladocerans, chironomids, heteropterans, coleopterans, molluscs, amphibians, fish and macrophytes. Per sampling day, we were able to sample three neighboring ponds.

   Below are some of the sampling methods used:

   1. (ABIOTIC) For the chemical analysis of water quality and the study of bacterioplankton and phytoplankton, we used a tube sampler (length 1.5 m; diameter 75 mm) to collect a depth-integrated sample of pond water in the open water part of each pond.

   2. (BIOTIC) To obtain samples for bacterioplankton, we filtered two volumes of approximately 200 ml over a 0.2 μm MF-Millipore MCE filter, which were then stored on liquid nitrogen.

   3. (BIOTIC) For zooplankton, we used the tube sampler to collect 6-L water samples at 8 different locations in the pond, according to a predefined grid (4 samples in the littoral and 4 samples in the open water area). All 8 samples were pooled and samples for crustacean zooplankton were obtained by filtering 40 L through a 64-μm conical net. An additional 3 L was filtered through a 30-μm net to concentrate the rotifer zooplankton. Zooplankton samples were stored on formaldehyde.

   4. (BIOTIC) We sampled phytobenthos by collecting the upper centimeter of sediment using a Plexiglas corer (diameter 30 mm) at two locations in the center of each pond; both samples were pooled for analysis and frozen (-80°C). For zoobenthos (chironomids), we sampled the upper 15 centimeters of sediment at 8 undisturbed locations by means of a hand corer (diameter 52 mm).

   5. (BIOTIC) We collected semi-quantitative samples of macroinvertebrates (coleopterans, heteropterans, molluscs) by using a D-shaped sweep net (25 cm x 25 cm; mesh size: 250 μm) in the open water as well as in the submerged, floating and emergent vegetation. The net was hauled through the entire water column for a total of two minutes, representing each of the habitat types in proportion to its estimated cover.

   6. (BIOTIC) We assessed fish communities applying the technique of point abundance sampling with electrofishing. In each pond the anode was immersed at eight

haphazardly chosen locations and fish were collected with a hand net.

7. (BIOTIC) Amphibia were assessed in the field from sweep net samples (diameter mouth 45 cm; mesh size: 1 mm) and by visual inspection of aquatic macrophyte stands and pond margins.

8. (ABIOTIC) We analyzed water samples for the concentration of chlorophyll a, nutrients (total phosphorus and nitrates), alkalinity and some major ions (calcium, chloride and sulphate ions, water hardness). Sulphates, chlorides, calcium, alkalinity and hardness were measured following standard methods according to the Hach Water Analysis Handbook (HACH, 1992).

Pre-workshop notes: this data is collected from a large project. I represents sampling events for both biotic and abiotic measurements. It serves as a valuable example, because it requires us to determine how much information should be modelled? Is it important to include sampling events for DNA sequencing, if those sequences aren't used in taxonomic identification/barcoding for example?

Workshop summary:
- *See **work group report***
- *They (Ana Cristina, Cardoso, Wouter Koch, Rui Figueira & Andre Heughebaert) felt it was possible to publish all the data in sample event format in a single dataset.*
- *They chose to represent the sampling at a pond as a parent event with abiotic measurements, and sampling for species groups as child events with their own measurements, etc.*
- *They explained the raw cross table data would have to be transformed manually or using custom scripts, and that some information about the event would have to be derived from the locality itself.*
- *They used the DwC terms, and GBIF Madagascar training event resources, saying they were sufficient*

Post-workshop follow up:
- *Andre has converted the data derived from amphibians sampling events into a sample event dataset, and Kyle has sent him suggestions on how to improve it further. The set of scripts created by Andre will be put in GitHub for the benefit of others.*
- *[DOCUMENTATION] They recommended making guidelines for how to deal with/represent large projects either as single datasets, or breaking them up into many constituent datasets. Breaking a project up into multiple datasets means metadata will get repeated often, requiring extra effort. It would also require good cross linking between the datasets.*
- *[DOCUMENTATION] They recommended making guidelines for publishers on how to use cross tables, for example by developing a standard cross table formatted template for publishers to use.*
- *[DOCUMENTATION] They recommended making guidelines for how to migrate existing occurrence datasets into sample event format. The guidelines should address the*

following questions:

- *should the sample event version replace the existing occurrence version, or should both versions be kept online at the same time?*
- *If replacing, should the new sample event version be assigned a brand new DOI?*
- *What are the benefits of going to the trouble of converting the dataset to use the sample event format?*

- *[VISUALISATION] They recommended GBIF.org show datasets that are related under the same project for example (e.g. using the Project.id field in EML currently not parsed by GBIF).*
- *[VISUALISATION] They recommended GBIF.org show absence data, as long as its shown in a way that doesn't confuse users. In this dataset, dry ponds/pools would have no measurements or occurrences, but would still be valuable to show.*
- *[DOCUMENTATION] They recommended the entire process of publication be documented. For example, the scripts for transforming and validating the data could be stored in GitHub.*
- *[DOCUMENTATION] They recommended issues related to the published dataset could be managed in GitHub. This is what INBO is currently doing for their datasets.*
- *[STANDARDS] They recommended that for sensitive/protected species, possible solutions include:*
  - *Simply removing these species from the dataset*
  - *Publishing the species identifications at Genus level only*
  - *Publishing the sensitive/protected species in a separate dataset*

Sophie Pamerlon (GBIF France)

1. Carab beetles. Original Excel data provided, but it must remain private. This is a sample-dataset about carab beetles (Coleoptera) that's been recorded between 2007 and 2012 in Brittany (near the city of Rennes) by using **Barber traps**. The "abundance" field represents the number of individuals from each species that has been recorded in each Barber trap during each sampling session.

   Pre-workshop notes: this dataset is missing spatial and temporal information for each sampling event. The identifications are missing for some species. More information about the sampling protocol/traps needed still. This dataset can serve to demonstrate the type of data received from publishers, and can help us understand what required fields and metadata should be asked for to make the data publishable.

   ***Workshop summary**:*

   - ***Were you able to find out enough information (metadata) about the dataset to***

understand it? Could you fill in the missing columns in the Overview - Example Datasets spreadsheet?

Yes but it was sometimes difficult to interpret because of the lack of metadata and explanations about field names from the original spreadsheet (e.g. session_code or eltp fields that had no unit/were not documented). We were able to guess and fill in some missing mandatory terms such as eventDate (we extracted the year from the session_code field), basisOfRecord (we chose « Occurrence » because it couldn't be any other option) and countryCode (we checked the verbatim coordinates with QGIS).

- ***Which issues (topics below) did you encounter while working on the example dataset?***
    - ***What format (event vs. occurrence) to use, also which extensions***
        - *The choice of the DwC core was not difficult given that it was sample-based data collected using traps that were inspected several times a year. We also used the Occurrence core as an extension in order to map the relevant fields such as scientificName, basisOfRecord, organismQuantity and organismQuantityType*
    - ***How to handle presence / absence data***
        - *The dataset we used didn't contain absence data or any other information indicating absence.*
    - ***How to record the sampling protocol, abundance, other DwC fields***
        - *The type of trap was given in the original dataset (Barber traps). Abundance was also known but it was not very clear (a single field with a numeric value without units). Other fields were also given or could be extracted and/or transformed (eventDate, decimalLatitude and decimalLongitude).*
- ***How to transform the raw data into DwC tabular form***

    This was not an issue given that the original dataset was already an Excel spreadsheet. We just changed the names of most of the fields in order to facilitate the Darwin Core mapping (some fields were not exploitable such as eltp that contained information that couldn't be identified).
- ***For each issue you encountered, did you find a solution / recommendation? Please descrIbe.***

    Sample-size value : this information was not given so we had to assume it (for the sake of the exercise, and given that we couldn't contact the dataset provider, we chose the duration between the sampling events and assigned a value of 2 weeks according to our knowledge of such sampling protocol)

    The main issue was to identify each original field and its content and interpret the relationships between the fields.
- ***Did you successfully publish the example dataset into the sample event format?***

Yes, but it is private : http://eubon-ipt.gbif.org/resource?r=test_sophie_manash_maxim

- **What documentation did you use to learn more about how to publish sample event data? E.g. what DwC terms to populate, how to populate DwC terms, list of extensions available for use, etc.**

We used the Darwin Core Reference Guide : http://rs.tdwg.org/dwc/terms/index.htm

- **Was the documentation sufficient, or was there something missing? E.g do you need these documentation in your local language, or is English sufficient?**

The English documentation was sufficient for us but it would be great to translate it into other languages as well as it would be more helpful for new data publishers (for example, there are people in Russia and France interested in using the IPT to publish their data).

Post-workshop follow up:

- *Kyle has sent the group a set of suggestions on how to improve the dataset, with the ultimate goal of having it published through GBIF some day:*
    - *Add occurrenceID - otherwise occurrrence records cannot be indexed by GBIF*
    - *Used basisOfRecord "MaterialSample" instead of "Occurrence", which is ambiguous.*
    - *Use individualCount instead of organismQuantityType "individuals" to store original recording for entire sampling event*
    - *Indicate that you have used QuantumGIS to verify the location of the data using* http://rs.tdwg.org/dwc/terms/#georeferenceVerificationStatus
- *[DOCUMENTATION/STANDARDS]They (Sophie, Maxim, and Manash) recommended using the fieldNumber in the eventID to make it more readible (+1 Anders F)*
- *[DOCUMENTATION] They recommended IPT documentation be made available in French, Russian, etc. (although Dimitry B insists English should remain the predominant language of DwC terms, with focus instead on better definitions)*

Dimitri Brosens - INBO

1. InboVeg Niche Vlaanderen. Data in DwC data provided, but it must remain private. The original data is managed in https://www.inbo.be/en/inboveg based on the same datamodel as Turboveg. **Sampling methodology:** Vegetation plot surveys were performed between May and July 2002 near selected Piezometers (INBO-watina project). In forests the plot measures 10X10m² in open vegetations this plot measures 3X3 m². For every plot, all occurring species were determined and for every vertical layer (Trees, Shrubs, Herbs and Mosses).The protocol as identified in INBOveg is "classic" and for the estimation of the coverage we use the Londo Scale. By 'Classic' we mean that the total coverage can be more than 100%. Additional information about the sampling protocol can be found here:

https://github.com/LifeWatchINBO/data-publication/tree/master/datasets/inboveg-niche-vlaanderen-events

Pre-workshop notes: *this dataset uses releves, measured in Londo scale, not Braun-Blanquet scale. This is a quantity type not included in the GBIF Quantity Type vocabulary: http://rs.gbif.org/vocabulary/gbif/quantity_type_2015-07-10.xml Can its Releve data map to the GBIF Releve extension: http://rs.gbif.org/sandbox/extension/releve_2016-02-26.xml ? What format is the original data in (e.g. abundance matrix) and should the IPT be able to support this?*

Workshop summary:
- *Dimitry successfully created a DwC-A in sample event format using the latest GBIF Releve extension (sent to Kyle privately).*
- *Dimitry reported that the GBIF Releve extension matches Inboveg's database quite nicely, since both are based on Turboveg's database schema.*

Post-workshop follow up:
- *Kyle reviewed the DwC-A created by Dimitry and has provided him with some suggestions on how to improve the dataset:*
  - *Invalid use of parentEventID, since it doesn't reference an exisiting event with that ID*
  - *Invalid use of the Relevee extension: have many Releves to a single event. In theory, each Releve should represent a single sampling event. Plus there is very little variation between all these Releves. Presumably they were carried out for different vegetation layers, but this isn't indicated.*
  - *Invalid Relevee.coverScale (e.g. 3). This should be "Londo scale"*
  - *Would be nice to indicate what vegetation layer each occurrence was observed in using the sandbox GBIF Vegetation Layer vocabulary: http://rs.gbif.org/sandbox/vocabulary/gbif/vegetation_layer.xml*
- *[STANDARDS] Dimitry recommended adding Londo scale to the GBIF Quantity Type vocabulary.*
- *[STANDARDS] Dimitry recommended using MoF extension instead of custom extensions for handling user-defined layers for different sampling protocols such as Braun-blanquet/Relevees.*
- *[DOCUMENTATION] Dimitry recommended trying to alleviate confusion between DwC recommendations and IPT requirements. For example, BoR being required by IPT, but not in DwC.*
- *[STANDARDS] Pedro said researchers need to be able to get the relevee data out in the same way it comes in because they are very habitual scientists.*
- *[DOCUMENTATION] Pedro said it is helpful interpreting Relevee data, to see the original relevee sheet. Therefore a recommendation should be made to publishers to try and include a link in their data to the original relevee sheet.*

- *[DOCUMENTATION] Pedro said it is valuable to include links to published works related to the data, such as journal articles, project notes, thesis, etc. Certainly this helps in understanding the data better.*
- *[VISUALISATION] Pedro said it would be nice to visualise the location of releves 1x1 km UTM squares, just like SIVIM does (see screenshots below). SIVIM allows users to search for relevees matching vegetation communities (syntaxons).*

Hanna Koivula (GBIF Finland)

1. Lepidoptera collection of Hannu Saarenmaa. Data publicly available on GBIF Finland IPT: http://ipt.digitarium.fi/resource?r=semf **Sampling methodology**: mainly consists of about 10-20 different quantitative moth monitoring schemes which Hannu ran:

    1. Rovaniemi city 1980-85
    2. Vanttauskoski 1981-1984
    3. Pallasjärvi 1981-1984
    4. Ruovesi Susimäki 1992-1993
    5. Koli 1994-1995
    6. Eno Kirjovaara 1997
    7. Kungsmarken 1999-2005 Sweden
    8. Magleby skov 2005-2006 Denmark
    9. Ilomantsi Pappilanvaara 2008-2015
    10. Ilomantsi Kivivaara 2009
    11. Ilomantsi Möhkö 2008-2009
    12. Ilomantsi Möhkönvaara 2010-2011
    13. Värtsilä 2014-2015

    Pre-workshop notes: This is a great example of a dataset illustrating various issues.

    1. It uses multiple sampling protocols for the same group (order) of species. Therefore this presents a challenge of unifying the abundance and coverage statistics for example.
    2. It uses sample size in trap days. While it says "days" it is really "nights". It might be more meaningful to use "hours" instead, which would properly take into account the fact that here in the north we have very short nights and consequently poor catches. For instance, if the length of night is 6 hours, and it runs for a week, we have effective catching time of 42 hours. Often the lamp goes on while there still is dusk and it is not attracting until one hour later. Hannu thinks he should really use sunset/sunup times

with some margin when sun is a few degrees below the horizon (so called nautical night). But Hannu is not aware of any standard in moth monitoring schemes for how to represent this.

3. For 2008 onwards, you will find in the data separate observations and related (contained) specimens which may have been retained in my collection for evidence. Those specimens have a pointer to their mother occurrence records in the field associatedOccurrences. Hannu has been struggling with this in the star schema, because it creates duplicate records. We'd really need a snowflake schema.

4. Absence can be inferred depending on whether species is part of Macrolepidoptera or Microlepidoptera: All Macrolepidoptera have been identified and counted - absence can be inferred from lack of record. Microlepidoptera have been identified where possible, but their numbers are not necessarily quantitative, and lack of record does not mean absence.

Workshop summary:

1. Sample-based data publication; reflections on semantics and logic

The relation between <u>events</u>, <u>occurrence data</u> and <u>sample-based data</u>

- Events and occurrences may be published with separate datasets. However, forming the base of an event – that may or may not have a mixed basis of records that each require individual identifiers - basically all records are observations that can have documentation or "evidence" in the form of specimens, illustrations, multimedia files, sequences or other measurements.
- **Occurrences may be the discovery system**, hence GUIDs for occurrences that are linked to events are necessary but we also need a pathway "back" to the records metadata/the event – thus in addition to occurrences identifiers we need **categories in controlled vocabularies for events to make them searchable**.
- Original/first observation tracking data, bird ringing or other continuous monitoring of possibly the same individuals should have identifier that stays with individuals "forever" – but is that individual then the event?

2. Sample-based data publication; reflections on presence-absence data

<u>Presence</u>–<u>absence</u> data

- Absence data should be **published with a check-list** for the relevant location and time, which however unfortunately must leave out those species that would have presumably been observed, but are not part of the "normal" species composition.
- For taxonomic and biogeographical/ecological reasons such checklists will have to be dated or provided solely alongside the sampling event in question.
- Species that turn up far outside of their home range (or natural habitat), and which do not coincide with local check-lists (e.g. invasives) are challenging; never possible to contain in a monitoring scheme with check-list unless the latter

is supplemented by "known" expected species.

3. Sample-based data publication; reflections on "new" data types

"New" data types require special attention

- Tracking data, bird ringing and continuous monitoring (of possibly the same individuals) will require identifiers that stay with individuals since the first /original observation, but is that individual then an event, an observation or a specimen?
- Such events/records should, when possible, be followed by repeated monitoring of conditions and effects
- Live collections represent a kind of "monitoring" data.
- Controlled vocabulary needed for sampling methods, which should be searchable in the metadata or event to be able to explain the fitness-for-use for the event/dataset.
- For sampling methods that measure activity rather than quantity, absence is a relative term and needs to be expressed with probability.
- Not all datatypes need to be in DwC, but links must be established to (all kinds of) relevant databases that relate to occurrence/check-list (including barcodes). Many relevant "things" may be expressed by different standards but need not be contained in DwC.
- Special attention should be paid to social media big data extracts. How?
- Verbatim description is still needed but to be made searchable event metadata would need controlled vocabularies (e.g. sampling protocols of other "categories")

Post-workshop follow up:

- [DOCUMENTATION] They recommend guidelines explaining the relations between the events and occurrences (need to check if the primer does a good job of this).
- [STANDARDS/DOCUMENTATION] They recommended that a checklist with timestamp/version be published together with the sample event dataset, in order to properly represent the species that could be observed at the time and place of sampling. This would allow for accurate presence/absence data being recorded. If possible, the checklist should list and categorise known/expected/normal species composition plus the unexpected/invasive species. For taxonomic and biogeographical/ecological reasons, this checklist would exist solely within the context of the sample event dataset. An investigation is needed how best to publish both datasets together.
- [STANDARDS] They recommended a controlled vocabulary of sampling protocols divided into categories, in order to properly enter sampling protocols, and explain the fitness-for-use of a dataset. The verbatim sampling protocol would remain searchable, but such a vocabulary would allow for easier browsing and integration of the sampling data.
- [STANDARDS] They highlighted that some sampling methods measure activity rather than quantity, absence is a relative term and needs to be expressed with probability.

This may require new DwC terms to record the activity and its probability in %chance for example.

- [DOCUMENTATION] They recommend guidelines for continuous monitoring of the same individual, whereby live collections represent a kind of "monitoring" data:
  - It OK to recommend http://rs.tdwg.org/dwc/terms/#organismID as the appropriate DwC term to store an ID for the individual being tracked.
  - In sample event format, a single event record could exist for each individual being tracked, with associated occurrences where it was tracked/recorded.

Other recommendations and follow ups:

- [DOCUMENTATION] Produce an updated version of the Primer originally made by Eamonn: http://www.gbif.org/sites/default/files/gbif_IPT-sample-data-primer_en.pdf
  - For each DwC term, have better documentation for how to fill them in. E.g., construct eventID from fieldNumber so that it is more human readable. What is the intended meaning of each term?
  - Keep the Primer online, instead of in a PDF
  - Should include some basic level documentation on the benefits of standardisation, widely underappreciated
- [DOCUMENTATION] Include guidelines in documentation for trawl tracking using a LINESTRING WKT shape with start and end point allow the direction of the trawl to also identified
- The following bugs and enhancements in GBIF.org were identified during the workshop:
  - Bug on GBIF.org dataset page with zoom level on bounding box: http://dev.gbif.org/issues/browse/POR-3086
  - Add first and last occurrence request to GBIF Occurrence API: http://dev.gbif.org/issues/browse/POR-3093
  - Map occurrence search results: separate record types (BoR) by color: http://dev.gbif.org/issues/browse/POR-3083
  - Index Event records: http://dev.gbif.org/issues/browse/POR-3095
  - Index Measurement or Facts records: http://dev.gbif.org/issues/browse/POR-3096
  - Visualise Event records: http://dev.gbif.org/issues/browse/POR-3094
- [STANDARDS] Kyle needs to follow up with Nabil about generating a metagenomics example dataset in DwC-A sample event format.
- [MOBILISATION] Siro needs to follow up with Liam regarding mobilising EIA data from Shell.
- [MOBILISATION] Kyle needs to follow up with Peter Desmet about converting the INBO Bird Tracking dataset into sample event format.
- [MOBILISATION] A user of TV3's export-to-DwC-A-in-sample-event-format feature requested to present at EVS 2017, marketing this new feature to TV's primary user base.
- [MOBILISATION] Nodes are requested to approach thematic groups doing standardised sampling (e.g. Freshwater monitoring/BioFresh, EU funded projects) to help mobilise sample event datasets through the GBIF network.

- [MOBILISATION] Nodes are requested to assist collecting a variety of sampling protocols in common use, with the aim of producing a sampling protocol vocabulary divided into categories. E.g. each EU funded project likely has its own set of protocols. This work is something that could be funded under COST.
- [DOCUMENTATION] Nodes such as Norway that are advanced in the use of GUIDs, are requested to build a set of recommendations on how to implement GUIDs at a Node or institution in order to help pave the way for other Nodes for example.
- [DOCUMENTATION] Nodes are requested to lead publication of a scientific paper, which serves as a much more solid reference than a webpage.

Darwin Core data types:

http://www.semantic-web-journal.net/system/files/swj995.pdf
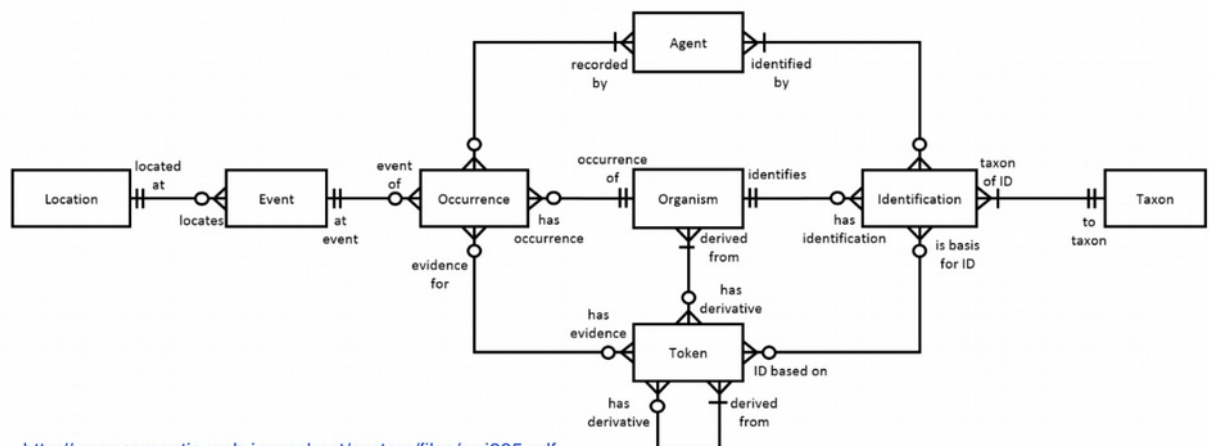
Fig. 1. Entity-relationship diagram of the Darwin-SW model using crow's foot notation with classes and relationships described in English text

http://www.semantic-web-journal.net/system/files/swj995.pdf

http://www.semantic-web-journal.net/content/darwin-sw-darwin-core-based-terms-expressing-biodiversity-data-rdf-0

SIVIM

Sistema de Información de la
Vegetación Ibérica y Macaronésica
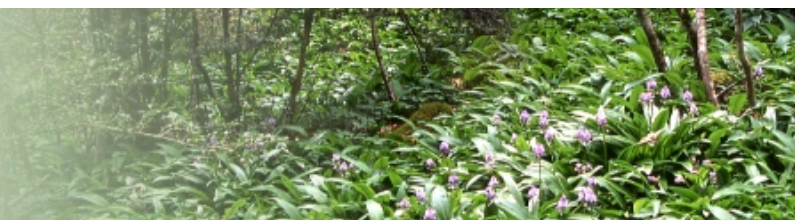
‹ **Área de distribución de la comunidad:** *Notobasio syriacae-Scolymetum maculati* **Rivas Goday ex Ladero, Socorro, Molero, M. López, Zafra, Marín, Hurtado & Pérez-Raya 1981 subass.** *carduncelletosum cuatrecasii* **García-Fuentes 1996** ›

Ecología   Sinónimos   Bibliografía   Composición   Mapa   gMapa

Opciones
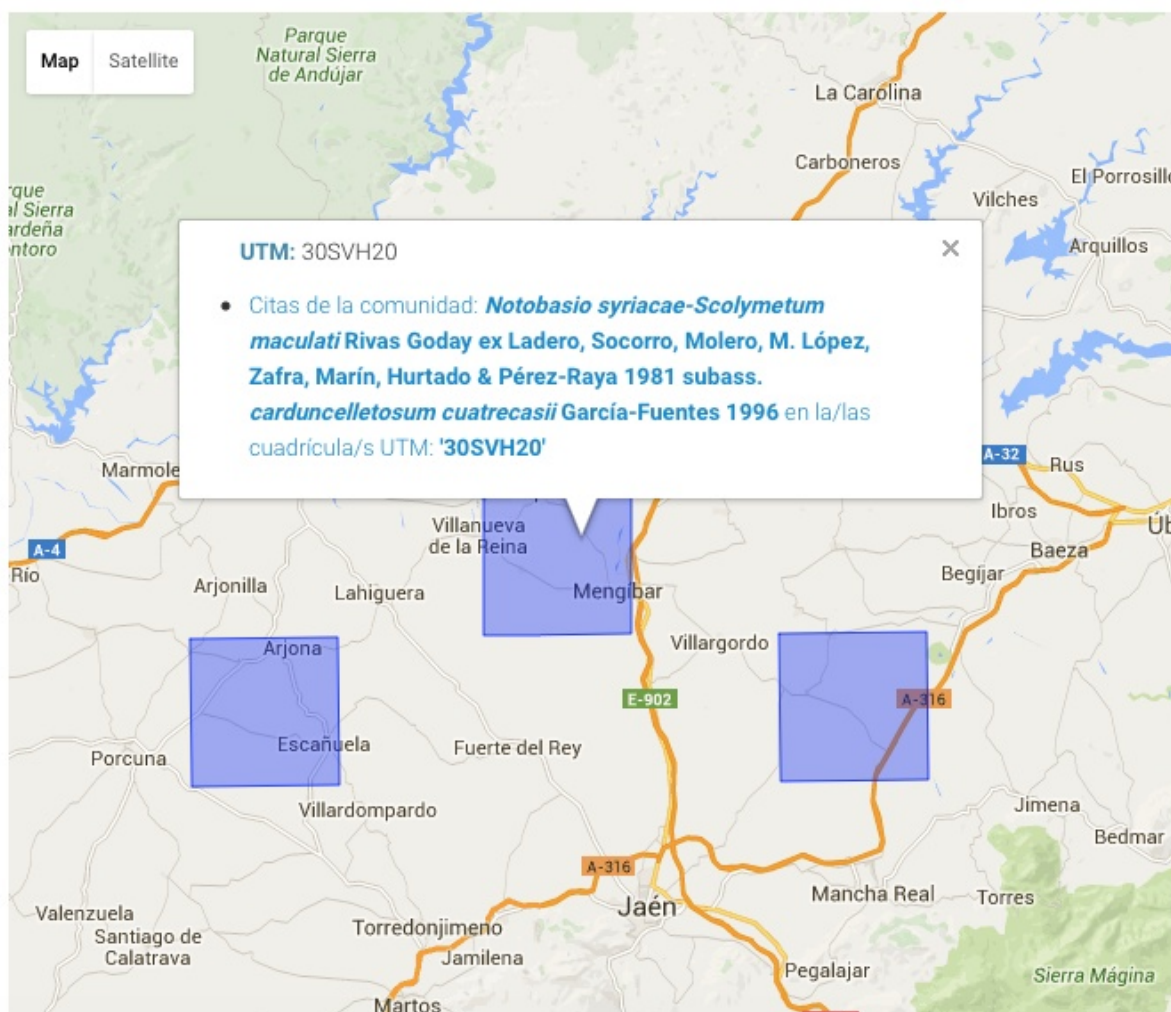
UTM: 30SVH20

- Citas de la comunidad: *Notobasio syriacae-Scolymetum maculati* Rivas Goday ex Ladero, Socorro, Molero, M. López, Zafra, Marín, Hurtado & Pérez-Raya 1981 subass. *carduncelletosum cuatrecasii* García-Fuentes 1996 en la/las cuadrícula/s UTM: **'30SVH20'**

García Fuentes, A.; (1996); Vegetación y Flórula del Alto Valle del Guadalquivir: Modelos de Regeneración.; Tesis Doctoral. Universidad de Jaén.; tabla 081

	1
Notobasis syriaca (L.) Cass. in Cuvier	+
Silybum marianum (L.) Gaertn.	+
Carduncellus cuatrecasassi G. López [sub]	1
Dactylis glomerata L.	1
Pallenis spinosa (L.) Cass.	+
Dittrichia viscosa (L.) Greuter	+
Imperata cylindrica (L.) Raeusch.	+

inventario: 1

orden del inventario en la tabla: 5
código de inventario: S-P21003
superficie inventariada: 50 m^2
localidad: Proximidades a Mengíbar
altitud: 300 m
utm: 30SVH20
Sintaxon: *Notobasio syriacae-Scolymetum maculati* Rivas Goday ex Ladero, Socorro, Molero, M. López, Zafra, Marín, Hurtado & Pérez-Raya 1981
sub: *Notobasio syriacae-Scolymetum maculati* Rivas Goday ex Ladero, Socorro, Molero, M. López, Zafra, Marín, Hurtado & Pérez-Raya 1981 subass. *carduncelletosum cuatrecasii* García-Fuentes 1996

En estos inventarios pueden faltar algunos táxones \\n originariamente determinados de manera incierta

Exporta XML ◀

Visualización de la tabla: García Fuentes, A. 1996 Vegetación y Flórula del Alto Valle del Guadalquivir: Modelos de Regeneración. Tesis Doctoral. Universidad de Jaén. T.081

García Fuentes, A.; (1996); Vegetación y Flórula del Alto Valle del Guadalquivir: Modelos de Regeneración.; Tesis Doctoral. Universidad de Jaén.

	1	2	3	4	5	6	7
Notobasis syriaca (L.) Cass. in Cuvier	3.4	.	2.2	1.1	+	1.1	.
Carthamus lanatus L.	.	.	3.4	1.1	.	.	1.1
Onopordum nervosum Boiss.	.	.	4.4	.	.	.	.
Silybum marianum (L.) Gaertn.	1.1	1.1	+	.	+	.	.
Echinops strigosus L.	.	.	+	.	.	1.2	.
Carduus bourgeanus Boiss. & Reut.	.	1.1	+	.	.	.	1.1
Carlina corymbosa L.	.	.	.	1.1	.	.	2.2
Carduncellus cuatrecasassi G. López [sub]	.	.	.	2.3	1.1	3.3	2.2
Malva multiflora (Cav.) Soldano, Banfi & Galasso [sub]	2.2	.	2.2	.	.	.	.
Avena sterilis L.	2.2	.	1.1	+	.	+	.
Hordeum murinum L. subsp. *leporinum* (Link) Arcang. [sub]	2.2	2.2	.	.	.	.	.
Avena X_haussknechtii	.	2.3	.	.	.	.	.
Vicia lutea L. subsp. *vestita* (Boiss.) Rouy	.	+	1.1	.	.	.	.
Lolium rigidum Gaudin [sub]	.	+	1.1	.	.	1.1	.
Phalaris minor Retz.	.	.	+	.	.	.	.
Hirschfeldia incana (L.) Lagr.-Foss.	.	.	+	.	.	.	.
Crepis vesicaria subsp. *haenselieri*	.	.	+	.	.	.	.
Scabiosa simplex Desf.	.	.	.	+	.	.	.
Centaurea melitensis L.	.	.	.	+	.	.	+
Phlomis lychnitis L.	.	.	.	+	.	.	+
Dactylis glomerata L.	.	.	.	.	1.1	+	.
Scorzonera laciniata L.	.	.	.	.	.	1.1	.
Catananche lutea L.	.	.	.	.	.	1.1	.
Brachypodium retusum (Pers.) P. Beauv.	.	.	.	.	.	.	+
Ecballium elaterium (L.) A. Rich. in Bory	2.3	.	.	.	.	.	.
Sinapis alba L.	2.2						

Annex III.
Workshop on sampling event data - Summary of recommendations from workshop

Summary of recommendations

Workshop on sampling event data, Lisbon 18-19 April 2016

Prepared by Kyle Braak, Dag Edersen and the workshop participants

Most updated version available at https://goo.gl/XwyS5A

Introduction

This is a summary of recommendations on how GBIF can improve the documentation, standardisation, visualisation, and mobilisation of sampling event data collected during the workshop on sampling event data held in conjunction with the 8th European GBIF Nodes Meeting in Lisbon, 18-19 April 2016.

Note this page is also accessible at the following short URL: https://goo.gl/XwyS5A. Comments are welcome.

Documentation

During the workshop participants used the following resources while working on their sampling event datasets:

- IPT Sample Data Primer
- DwC Terms Quick Reference Guide
- Materials from GBIF Nodes publishing training event at GB22 Madagascar 2015
- GBIF Resources (rs.gbif.org)
- IPT How To Publish Guide

Participants discovered that the Primer was out of date, and failed to answer lots of questions related to publishing sampling event data. Based on feedback collected during the workshop, it is suggested that the IPT Sample Data Primer **be brought up-to-date** and incorporate the following set of improvements:

- Advise publishers to learn as much as possible about the sampling event, especially the sampling methodologies, before attempting to try and standardise it into DwC.
- Better guide users on how to fill in DwC terms related to sampling event data.
  - dwc:eventID should be a persistent globally unique identifier. Remember to reuse existing stable identifiers. Do not create a new identifier for the event when one already is declared.
  - In the absence of a GUID, and as a last resort, reuse the original fieldNumber.
- Guide users how to publish presence/absence data in DwC with the proper use of

dwc:occurrenceStatus
- Better guide users how to publish a hierarchy of events (recursive data type) with the proper use of dwc:parentEventID
- Advise publishers to use GUIDs, coupled with guidance on how to create GUIDs for applicable fields such as dwc:occurrenceID, dwc:eventID, dwc:organismID and dwc:locationID. For example, it is possible to use http://www.geonames.org/ to find (or even generate new) identifiers for dwc:locationID, e.g. http://sws.geonames.org/10793757/ is a GUID for a lake in Greenland.
- Explain what DwC terms are required by IPT in order to publish, and what terms are simply considered required in terms of completeness.
- Guide users how to obfuscate the location of sensitive species, such as by:
    - *Simply removing these species from the dataset*
    - *Publishing the species identifications at Genus level only*
    - *Publishing the sensitive/protected species in a separate dataset*
    - *Publish obfuscated sensitive data points in the main dataset and publish non-obfuscated details in an access-limited separate dataset, both datasets including all data records*
- Guide users how to enter verbatim descriptions. For example, the ID or code given to the original event should be entered into dwc:fieldNumber; the ID or code given to the original occurrence observation should be entered into dwc:recordNumber.
- Better explain the relations between the events and occurrences.
- Provide recommendations on how to work around limitations of DwC-A star schema such as not being able to relate measurements and facts to both events and occurrences in the same dataset. The current work around requires publishers to publish separate datasets. Note OBIS is prototyping an Extended Measurement or Facts Extension that could also help overcome this limitation. Discussion on this prototype extension is taking place in GitHub here. However, issues raised that this prototype extension does not explicitly make it clear if the measurement or fact relates to an occurrence or an event. One alternative is to add resourceID (and perhaps resourceType?) instead of adding eventID (and occurrenceID) as attribute to the measurement or fact extension as is explored by the OBIS extension.
- Provide a recommendation on how to publish data produced from large projects. The current recommendation is to publish a single dataset, because dividing it into multiple datasets results in more duplication of effort entering metadata. Publishers insisting on publishing multiple datasets should link them using Project.ID in EML.
- Advise publishers on how to link related datasets that come out of the same research context so they can be easily retrieved by the users. Publishers may have to publish separate datasets in order to work around the limitations of the DwC-A star schema. Publishers may also choose to publish separate occurrence datasets derived from the same sampling events. The current recommendation is to link them using Project.ID.
- Provide rationale and guidance for migrating existing occurrence datasets to sampling event format. The following questions need to be answered:
    - Should the sampling event version replace the existing occurrence version, or should

   both versions be kept online at the same time?
   • If replacing, should the new sampling event version be assigned a brand new DOI?
   • What are the benefits of producing the sampling event version?

- Provide recommendations on how to make the dataset easier to interpret by including links to related published works such as journal articles, project notes, thesis, etc.
- Provide recommendations on how include supplementary media in order to make the data easier to interpret. For example for vegetation data, it is helpful to include a link to the original scanned releve sheet when interpreting the data.
- Provide recommendations on how to represent the sampling area by choosing the appropriate WKT shape or simple latitude/longitude point location. Done correctly, the direction sampling was carried out can also be derived. For example, an ocean trawl line represented using a WKT shape LINESTRING allows the direction of the trawl to be determined based on the standard notation for writing the start and end points.
- Provide a recommendation on how to manage issues related to the dataset using GitHub's issue management system, just like INBO does for example.
- Provide a recommendation on how to make custom scripts and programs (e.g. for transforming cross table data) publicly available using GitHub, for the benefit of other publishers, just like INBO does for example. The recommendation should encourage users to include a detailed set of instructions on how to run the scripts to make them more usable.
- Provide a recommendation on how to model continuous monitoring of live individuals, such as bird tracking data by using dwc:organismID to store the ID of the individual being tracked and by using a single event for representing each individual being tracked (with associated occurrences where it was recorded).
- Translate it into other languages such as French, Russian, etc.

The [IPT How To Publish Guide](#) makes strong recommendations on what terms are required or highly recommended, plus a set of Excel templates to assist users format their data correctly. Based on feedback collected from the workshop, the following set of changes have been suggested to improve this guide:

- Remove dwc:countryCode as a required term for occurrences and events because country borders shift. Stronger emphasis should be put on recommending the exact coordinates instead.
- Remove dwc:sampleSizeValue/sampleSizeUnit as required terms for events. This allows linking species observations to a given sampling event without knowledge of the sampling intensity (e.g. information lost in translation for old datasets). It also allows the DwC sampling event format to still be used for non-quantitative data sets (.e.g presence/absence only).
- Explicitly require dwc:footprintSRS when providing dwc:footprintWKT (dwc:geodeticDatum relates exclusively to dwc:decimalLatitude & dwc:decimalLongitude).
- Regarding the Excel templates: advise publishers on how to fill in redundant terms appearing in both the event core, and occurrence extension. In general, the location of occurrences are inherited from the location of the event, unless more specific locations were recorded for individual occurrences.

- Advise publishers how to enter lists in DwC using the pipe ("|") character, for applicable terms.

Lastly, it was suggested that authors from GBIF and the wider biodiversity community should **publish a scientific article** explaining how sampling event data can now be standardised and published through GBIF.org. This will serve as a more authoritative reference than the Primer.

Standards

During the workshop participants attempted to transform or shoehorn their raw data into DwC sampling event format using available DwC terms. Participants discovered the DwC standard was inadequate for fully representing their data. Based on feedback collected during the workshop, the following changes to DwC and the current set of extensions and vocabularies has been proposed:

- Add new term "siteID": https://github.com/tdwg/dwc/issues/126
- Add new term "siteTreatment": https://github.com/tdwg/dwc/issues/128
- Add new term "samplingTaxaRange": https://github.com/tdwg/dwc/issues/127
- Add new term "layer": https://github.com/tdwg/dwc/issues/125
- Add new quantity type "Londo Scale" to the GBIF Quantity Type Vocabulary. Note the Londo decimal scale is an improvement to Braun-Blanquet scale, and more information about the protocol can be found here.
- Regarding creating a controlled vocabulary for dwc:samplingProtocol:
    - **Short term:** try to identify sampling protocols that are the same or comparable, by grouping occurrence records (associated to the event) by quantity type and taxa. For example, you could safely identify records coming from a "butterfly monitoring scheme", as long as a) its order is "Lepidoptera", b) its quantity is measured in "number of individuals", and c) it comes from a valid sampling event (e.g. with defined sample size in m2).
    - **Long term:** try to develop a hierarchical vocabulary, perhaps through a TDWG working group. The protocols could be divided into a) surveys vs sporadic observations, b) quantitative vs non-quantitative, c) terrestrial vs aquatic, etc.
- It was highlighted that some sampling methods measure activity rather than quantity. This may require a new DwC terms similar to organismQuantity & organismQuantityType in order to record the activity and its probability in percent chance for example. Further investigation is still needed.

Instead of modifying the IPT to handle cross tab table data, it was recommended we try to **develop a standard cross tab table template** to help users standardise how they enter data. This will allow them to take advantage of Excel's ability to flatten cross tab table data, and other custom scripts used to transform and validate cross table data.

There were two groups at the workshop that looked at vegetation data and **evaluated the use of the GBIF Relevé Extension** that enables relevés to be recorded in a standardised way based on Turboveg model. Fear was expressed, however, that this would lead to custom extensions having to be created for various sampling protocols. Alternatively, it was recommended that the Darwin Core Measurement or Facts Extension could be used for entering user-defined layers for various sampling protocols instead. More prototyping is needed to investigate if this should become recommended best practice.

One group at the workshop looked at how to **properly record presence/absence data**. Their recommendation was to publish a timestamped checklist together with the sampling event dataset, which represents the species composition that could be observed at the time and place of sampling given the sampling protocol (and/or the taxonomic coverage of the study and the expertise of the personnel carrying out identification). This would allow for accurate presence/absence data being recorded. In addition to the normal (expected) species composition, the checklist could include invasive (unexpected) species. For taxonomic and biogeographical/ecological reasons, however, this checklist would exist solely within the context of the sampling event dataset. An investigation is needed how best to publish both datasets together since multiple DwC-As cannot be bundled together in the same .zip folder.

Visualisation

Participants were all asked to imagine how they would like to see their sampling event datasets visualised. Christian Svindseth (GBIF Norway) produced a prototype sampling event browser in the data.gbif.no portal to provide inspiration. Other websites capable of visualising sampling events were also tested, such as SIVIM, an online tool for managing and visualising Iberian and Macaronesian vegetation data. Based on feedback collected during the workshop, the following set of issues has been submitted to the GBIF issue management site:

- Index Event records: http://dev.gbif.org/issues/browse/POR-3095
- Index Measurement or Facts records: http://dev.gbif.org/issues/browse/POR-3096
- Search Event records: http://dev.gbif.org/issues/browse/POR-3122
- Visualise Event records: http://dev.gbif.org/issues/browse/POR-3094

Additionally, participants helped collect the following set of bugs and enhancements submitted to the GBIF issue management site:

- Bug: GBIF.org dataset page has wrong zoom level on bounding box when there are no occurrences http://dev.gbif.org/issues/browse/POR-3086
- Bug: Skip absence data on indexing: http://dev.gbif.org/issues/browse/POR-2864
- Enhancement: Map occurrence search results: separate record types (BoR) by color: http://dev.gbif.org/issues/browse/POR-3083

- Enhancement: Make Project detail page showing its related datasets: http://dev.gbif.org/issues/browse/POR-3100
- Enhancement: Add first and last occurrence request to GBIF Occurrence API: http://dev.gbif.org/issues/browse/POR-3093

Mobilisation

The workshop and meeting were good opportunities to pinpoint new sources of sampling event data that can be mobilised in Europe such as environmental impact assessment (EIA) data. Nodes are kindly requested to approach thematic groups or EU funded projects doing standardised sampling such as BioFresh in order to help mobilise sampling event datasets. Nodes are also kindly requested to assist in identifying sampling protocols that are in common use, with the aim of producing a controlled vocabulary for sampling protocols (see above). Each EU funded project likely has its own set of protocols for example.

Existing occurrence datasets should be republished as sampling event datasets because this richer format maximises the usefulness of the data for users, and makes the dataset easier to interpret. For example, INBO has been requested to convert their Bird Tracking dataset into sampling event format, see: https://github.com/inbo/data-publication/issues/111

Lastly, Kyle Braak has kindly asked Nodes to identify users of Turboveg software within their network, and educate them it is possible to export vegetation data in DwC-A format using Turboveg's built-in DwC-A export feature (see Section 7.24 in the Turboveg User Manual). Kyle is currently working with Turboveg Lead Stephan Hennekens to update this DwC-A export feature to use the new sampling event format. Hopefully one Turboveg user can stand up, and present at the European Vegetation Survey conference in 2017 to advertise this new feature more widely, and ultimately help mobilise more vegetation data through GBIF.org.

Annex IV.
List of Participants

Participants to the 8th European GBIF Nodes Meeting

Serial	Name	Institution	Country
1	Rui Figueira	GBIF Portugal (organiser)	Portugal
2	Andre Heughebaert	Belgian Biodiversity Platform	Belgium
3	Liam Lysaght	National Biodiversity Data Centre, Ireland	Ireland
4	Wouter Addink	Naturalis Biodiversity Center	Nederland
5	Anders Telenius	Swedish Museum of Natural History	Sweden
6	Dimitri Brosens	Belgian Biodiversity Platform	Belgium
7	Dag Endresen	GBIF Norway	Norway
8	Manash Shah	Swedish Museum of Natural History	Sweden
9	Anne-Sophie Archambeau	GBIF France	France
10	Filipa Alves	GBIF Portugal (organiser)	Portugal
11	Ofer Steinitz	Israel Nature and Parks Authority	Israel
12	Tania Walisch	Musée national d'histoire naturelle Luxembourg	Luxembourg
13	Cristina Villaverde	GBIF Spain	Spain
14	Hanna Koivula	Finnish Museum of Natural History	Finland
15	Wouter Koch	Norwegian Biodiversity Information Centre	Norway
16	Anders G. Finstad	NTNU University Museum	Norway
17	Roger Caritg	GBIF Andorra	Andorra
18	Maxim Shashkov	Laboratory of Ecosystem Modelling IPBPSS RAS	Russia
19	Nils Valland	Norwegian Biodiversity Information Centre (NBIC)	Norway
20	Jörg Holetschek	Botanical Garden & Botanic Museum Berlin	Germany
21	Alberto González-Talaván	GBIF Secretariat	N/a
22	Roman Khalikov	Zoological Institute RAS	Russia
23	Ana Cristina Cardoso	European Commission, Joint Research Centre	Italy
24	Kyle Braak	GBIF	Denmark

Participants to the Workshop on Sample-based data publication

Serial	Name	Institution	Country
1	Rui Figueira	GBIF Portugal	Portugal
2	Andre Heughebaert	Belgian Biodiversity Platform	Belgium
5	Anders Telenius	Swedish Museum of Natural History	Sweden
6	Dimitri Brosens	Belgian Biodiversity Platform	Belgium
7	Dag Endresen	GBIF Norway (organiser)	Norway
8	Manash Shah	Swedish Museum of Natural History	Sweden
9	Anne-Sophie Archambeau	GBIF France	France
10	Sophie Pamerlon	GBIF France	France
11	Nabil Youdjou	Royal Belgian Institute of Natural Sciences	Belgium
12	Filipa Alves	GBIF Portugal	Portugal
13	Tania Walisch	Musée national d'histoire naturelle Luxembourg	Luxembourg
14	Hanna Koivula	Finnish Museum of Natural History	Finland
15	Wouter Koch	Norwegian Biodiversity Information Centre	Norway
16	Anders G. Finstad	NTNU University Museum Centre (organiser)	Norway
17	Maxim Shashkov	Laboratory of Ecosystem Modelling IPBPSS RAS	Russia
18	Jörg Holetschek	Botanical Garden & Botanic Museum Berlin	Germany
19	Nils Valland	Norwegian Biodiversity Information Centre (NBIC)	Norway
20	Roman Khalikov	Zoological Institute RAS	Russia
21	Ana Cristina Cardoso	European Commission, Joint Research Centre	Italy
22	Kyle Braak	GBIF (organiser)	Denmark
23	Pedro Arsénio	Instituto Superior de Agronomia/Univ. de Lisboa	Portugal

Annex V.
Selection of pictures of the Meeting, Workshop, Social Events and Excursion

Meeting

The meeting group photo is available for download at

http://www.gbif.pt/sites/default/files/8th_EurNodesMeeting_group_big.JPG

Workshop

The workshop group photo is available for download at
http://www.gbif.pt/sites/default/files/Workshop_SamplBasedDataPubish_group_big.JPG

Social Events

Excursion

Annex VI.
Press releases

GBIF Portugal organiza 8ª Reunião de Nós Europeus do GBIF em Lisboa

O Nó Português do GBIF vai organizar, entre os próximos dias 19 e 21 de Abril, a 8ª Reunião de Nós Europeus do GBIF no Instituto Superior de Agronomia, em Lisboa.

Com a presença dos representantes dos nós nacionais ou organizações de dezasseis países europeus, a reunião contribuirá para desenvolver a colaboração europeia em temas como a implementação de portais nacionais de dados de biodiversidade, mobilização de dados de biodiversidade de instituições e de ciência cidadã, a colaboração com a Rede Europeia de Informação sobre espécies exóticas (EASIN), a colaboração do GBIF no âmbito das diretivas europeias e o uso dos dados de biodiversidade, entre outros.

"A reunião, que é aberta a todos os colaboradores dos Nós GBIF da Europa, tem como objetivo definir o contributo dos Nós Europeus para o programa de trabalhos e plano estratégico do GBIF internacional", refere Rui Figueira, Coordenador do Nó Português, que acrescenta "A troca de experiência entre os nós contribuirá ainda para o enriquecimento e melhoria da capacidade nacional no serviço aos investigadores e instituições nacionais".

Os trabalhos serão precedidos por um Workshop Técnico sobre publicação de dados de amostragens e de inventários que é suportada pelo GBIF apenas desde o ano passado, o que constitui um avanço significativo já que estes dados são mais ricos em informação do que os registos de presença/ausência tradicionalmente disponibilizados através da organização.

O programa do 8ª Reunião de Nós Europeus do GBIF inclui ainda dois eventos sociais e uma visita ao Jardim Botânico da Ajuda, terminando com uma Excursão ao Parque Natural da Arrábida,

Toda a informação sobre o evento está disponível no portal do Nó Português do GBIF, em http://www.gbif.pt.

NOTAS PARA OS EDITORES:

1. Sobre o GBIF

O Sistema Global de Informação sobre a Biodiversidade (GBIF) é uma organização intergovernamental criada em 2001 para facilitar a partilha e acesso, de forma livre e gratuita, de dados de biodiversidade. Actualmente, são 64 os países signatários do Memorando de Entendimento do GBIF, dos quais 38 são membros votantes do Conselho de Administração da organização, entre os quais se inclui Portugal.

Alguns números do GBIF:

- o GBIF fornece um ponto de acesso único (através do seu portal e dos seus serviços web) a centenas de milhões de registos, livremente partilhados por centenas de instituições em todo o mundo, constituindo o maior banco de dados de biodiversidade na Internet;

- os dados acessíveis através do GBIF dizem respeito a registos de mais de 1,6 milhões de espécies, tendo sido recolhidos ao longo de três séculos de exploração da história natural e incluem observações recentes de cidadãos cientistas, investigadores e programas de monitorização automatizados;

- Mais de 1.700 publicações em revistas científicas com revisão por pares [sistema de validação dos resultados da investigação científica] citaram o GBIF como uma fonte de dados em estudos que analisam desde os impactos das alterações climáticas, a disseminação de pragas e doenças e as áreas prioritárias para a conservação até à segurança alimentar. Cerca de um destes artigos é publicado por dia.

2. Sobre o Nó Português do GBIF

O Nó Português do GBIF foi criado no Instituto de Investigação Científica Tropical em 2013, por indicação da FCT, tendo transitado para o Instituto Superior de Agronomia em Setembro de 2015. A sua principal missão é a promoção da participação das instituições nacionais como publicadores de dados através do GBIF, e fomentar o uso da informação sobre biodiversidade pela comunidade nacional em investigação científica e usos pela sociedade.

Desde a sua criação, o Nó Português promoveu o aumento para onze o número de instituições nacionais publicadoras de dados através do GBIF, que contribuem com mais de 400 mil dados de ocorrência de espécies, não só para o território nacional, mas também para outros países, nomeadamente os países africanos de língua portuguesa. São ainda publicadas checklists de espécies, quer para Portugal continental, quer para as regiões autónomas. A este facto acresce, enquanto recurso disponível para a comunidade nacional, mais de um milhão de registos de biodiversidade para Portugal, publicado por instituições estrangeiras.

O Nó procura também promover, em colaboração com o Secretariado Internacional do GBIF, a participação de outros países da CPLP no GBIF. Uma das formas têm sido através da produção de versões portuguesas de manuais e ferramentas sobre informática para a biodiversidade. Para além disso, foi recentemente realizado em Angola um workshop sobre o GBIF, estando planeado para Maio deste ano o mesmo em Moçambique. Para além da promoção da publicação de dados, o Nó Português facilitará também, a breve prazo, o acesso à informação a nível nacional através da implementação do portal de dados de biodiversidade, que facilitará o acesso e análise da informação nacional.

3.   Sobre o Meeting de Nós Europeus do GBIF

Os Nós Europeus do GBIF reúnem-se anualmente para desenvolver programas de trabalho conjuntos ao nível europeu, assim como coordenar a interação com outras regiões do GBIF, nomeadamente a africana e sul-americana. A atividade dos Nós, partilhada nestas reuniões, constitui uma base importante das atividades desenvolvidas pela rede GBIF. A reunião proporciona também um excelente ambiente de troca de experiências, reforçando a capacidade do Nó nacional e facilitando o desenvolvimento de novas oportunidades de cooperação. A reunião de 2015 ocorreu em Paris (França), tendo a de 2014 tido lugar em Bruxelas (Bélgica).

Informática da Biodiversidade: Nós Europeus do GBIF discutem em Lisboa criação de currículo

Os Nós Europeus do GBIF - Sistema Global de Informação sobre a Biodiversidade estão, desde terça-feira, reunidos em Lisboa no âmbito do seu encontro anual, sendo a criação de um currículo de Informática da Biodiversidade um dos temas em discussão.

A Informática da Biodiversidade diz respeito à recolha, comparação, integração, análise, previsão e disseminação de dados sobre os recursos biológicos, sendo um campo científico emergente.

Por influenciar, direta ou indiretamente, várias áreas da ciência aplicada relacionada com o bem-estar humano, é urgente a capacitação nesta área, quer seja através da criação de cursos de formação em instituições de Ensino Superior ou de programas de formação profissional.

O GBIF, enquanto entidade responsável pelo desenvolvimento e gestão da maior base de dados de Biodiversidade na internet, pode contribuir de forma importante para este processo, propondo-se fazê-lo através do COST, uma iniciativa da UE que visa possibilitar o estabelecimento de redes de investigação promovendo avanços científicos marcantes que conduzam a novos conceitos e produtos.

A reunião do GBIF em Lisboa, que decorre até ao final do dia de hoje no Instituto Superior de Agronomia, tem como objectivo a concretização do programa de trabalhos dos Nós Europeus do GBIF.

Este irá contribuir com vários tópicos para o programa de trabalhos e plano estratégico do GBIF a nível global para além do currículo em informática da biodiversidade, tais como o desenvolvimento de portais nacionais de dados, a publicação de dados de inventários de biodiversidade, a mobilização de dados de espécies invasoras, o uso de ferramentas de qualidade de dados, entre outros.

Os resultados da reunião contribuirão ainda para o aumento da capacidade do GBIF Portugal no apoio à comunidade nacional de utilizadores e instituições participantes no GBIF.

Notas para os Editores:

1. Sobre o GBIF

O Sistema Global de Informação sobre a Biodiversidade (GBIF) é uma organização intergovernamental criada em 2001 para facilitar a partilha e acesso, de forma livre e gratuita, de dados de biodiversidade. Actualmente, são 64 os países signatários do Memorando de Entendimento do GBIF, dos quais 38 são membros votantes do Conselho de Administração da organização, entre os quais se inclui Portugal.

Alguns números do GBIF:

- o GBIF fornece um ponto de acesso único (através do seu portal e dos seus serviços web) a centenas de milhões de registos, livremente partilhados por centenas de instituições em todo o mundo, constituindo o maior banco de dados de biodiversidade na Internet;

- os dados acessíveis através do GBIF dizem respeito a registos de mais de 1,6 milhões de espécies, tendo sido recolhidos ao longo de três séculos de exploração da história natural e incluem observações recentes de cidadãos cientistas, investigadores e programas de monitorização automatizados;

- Mais de 1.700 publicações em revistas científicas com revisão por pares [sistema de validação dos resultados da investigação científica] citaram o GBIF como uma fonte de dados em estudos que analisam

desde os impactos das alterações climáticas, a disseminação de pragas e doenças e as áreas prioritárias para a conservação até à segurança alimentar. Cerca de um destes artigos é publicado por dia.

2. Sobre o Nó Português do GBIF

O Nó Português do GBIF foi criado no Instituto de Investigação Científica Tropical em 2013, por indicação da FCT, tendo transitado para o Instituto Superior de Agronomia em Setembro de 2015. A sua principal missão é a promoção da participação das instituições nacionais como publicadores de dados através do GBIF, e fomentar o uso da informação sobre biodiversidade pela comunidade nacional em investigação científica e usos pela sociedade.

Desde a sua criação, o Nó Português promoveu o aumento para onze o número de instituições nacionais publicadoras de dados através do GBIF, que contribuem com mais de 400 mil dados de ocorrência de espécies, não só para o território nacional, mas também para outros países, nomeadamente os países africanos de língua portuguesa. São ainda publicadas *checklists* de espécies, quer para Portugal continental, quer para as regiões autónomas. A este facto acresce, enquanto recurso disponível para a comunidade nacional, mais de um milhão de registos de biodiversidade para Portugal, publicado por instituições estrangeiras.

O Nó procura também promover, em colaboração com o Secretariado Internacional do GBIF, a participação de outros países da CPLP no GBIF. Uma das formas têm sido através da produção de versões portuguesas de manuais e ferramentas sobre informática para a biodiversidade. Para além disso, foi recentemente realizado em Angola um workshop sobre o GBIF, estando planeado para Maio deste ano o mesmo em Moçambique. Para além da promoção da publicação de dados, o Nó Português facilitará também, a breve prazo, o acesso à informação a nível nacional através da implementação do portal de dados de biodiversidade, que facilitará o acesso e análise da informação nacional.

3. Sobre o Meeting de Nós Europeus do GBIF

Os Nós Europeus do GBIF reúnem-se anualmente para desenvolver programas de trabalho conjuntos ao nível europeu, assim como coordenar a interação com outras regiões do GBIF, nomeadamente a africana e sul-americana. A atividade dos Nós, partilhada nestas reuniões, constitui uma base importante das atividades desenvolvidas pela rede GBIF. A reunião proporciona também um excelente ambiente de troca de experiências, reforçando a capacidade do Nó nacional e facilitando o desenvolvimento de novas oportunidades de cooperação. A reunião de 2015 ocorreu em Paris (França), tendo a de 2014 tido lugar em Bruxelas (Bélgica).

CONTACTO PARA INFORMAÇÕES ADICIONAIS:

Rui Figueira (Coordenador do Nó Português do GBIF): 918 599 118

Annex VII.
E-mail sent to Academic/Research Institutions, Scientific societies and NGOs and List of institutions/organizations

E-mail text

Exmos Srs,

Vimos por este meio solicitar a divulgação do 8° Encontro de Nós Europeus do GBIF (Comunicado de Imprensa abaixo) organizado pelo Nó Português junto da vossa comunidade por considerar que pode ser do seu interesse.

Recorda-se que o Sistema Global de Informação sobre a Biodiversidade (GBIF) é uma organização intergovernamental criada para facilitar a partilha e acesso, de forma livre e gratuita, de dados de biodiversidade, fornecendo um ponto de acesso único a centenas de milhões de registos, livremente partilhados por centenas de instituições em todo o mundo, constituindo o maior banco de dados de biodiversidade na Internet.

O Nó Português do GBIF foi criado em 2013, tendo como principal missão promover a participação das instituições nacionais como publicadores de dados através do GBIF e fomentar o uso da informação sobre biodiversidade pela comunidade nacional em investigação científica e pela sociedade.

Texto a divulgar:

GBIF Portugal organiza 8ª Reunião de Nós Europeus do GBIF em Lisboa

O Sistema Global de Informação sobre a Biodiversidade (GBIF) é uma organização intergovernamental criada para facilitar a partilha e acesso, de forma livre e gratuita, a dados de biodiversidade, fornecendo um ponto de acesso único a centenas de milhões de registos, livremente partilhados por centenas de instituições em todo o mundo, constituindo o maior banco de dados de biodiversidade na Internet.

O Nó Português, criado em 2013, tendo como principal missão promover a participação das instituições nacionais como publicadores de dados através do GBIF e fomentar o uso da informação sobre biodiversidade pela comunidade nacional em investigação científica e pela sociedade. No âmbito das suas actividades, o Nó Português vai organizar, entre os próximos dias 19 e 21 de Abril, a 8ª Reunião de Nós Europeus do GBIF no Instituto Superior de Agronomia, em Lisboa.

Com a presença dos representantes dos nós nacionais ou organizações de dezasseis países europeus, a reunião contribuirá para desenvolver a colaboração europeia em temas como a implementação de portais nacionais de dados de biodiversidade, mobilização de dados de biodiversidade de instituições e de ciência cidadã, a colaboração com a Rede Europeia de Informação sobre espécies exóticas (EASIN), a colaboração do GBIF no âmbito das diretivas europeias e o uso dos dados de biodiversidade, entre outros.

"A reunião, que é aberta a todos os colaboradores dos Nós GBIF da Europa, tem como objetivo definir o contributo dos Nós Europeus para o programa de trabalhos e plano estratégico do GBIF internacional", refere Rui Figueira, Coordenador do Nó Português, que acrescenta "A troca de experiência entre os nós contribuirá ainda para o enriquecimento e melhoria da capacidade nacional no serviço aos investigadores e instituições nacionais".

Os trabalhos serão precedidos por um Workshop Técnico sobre publicação de dados de

amostragens e de inventários que é suportada pelo GBIF apenas desde o ano passado, o que constitui um avanço significativo já que estes dados são mais ricos em informação do que os registos de presença/ausência tradicionalmente disponibilizados através da organização.

O programa do 8ª Reunião de Nós Europeus do GBIF inclui ainda dois eventos sociais e uma visita ao Jardim Botânico da Ajuda, terminando com uma Excursão ao Parque Natural da Arrábida,

Toda a informação sobre o evento está disponível no portal do Nó Português do GBIF, em http://www.gbif.pt.

List of e-mailed institutions/organizations

	Academic/Research Institutions
1	CIBIO - Centro de Investigação em Biodiversidade e Recursos Genéticos
2	Ce3C
3	Escola de Ciências da Universidade do Minho
4	Faculdade de Ciências da Universidade de Lisboa
5	Faculdade de Ciências da Universidade do Porto
6	Faculdade de Ciências e Tecnologia da Universidade de Coimbra
7	Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa
8	Jardim Botânico da Madeira
9	Jardim Botânico da Universidade de Coimbra
10	Jardim Botânico do Faial
11	Museu de História Natural da Universidade do Porto
12	Museu Nacional de História Natural e da Ciência
13	Universidade – Faculdade de Medicina Veterinária, Faculdade de Psicologia e Ciências da Vida
14	Universidade Aberta:
15	Universidade da Madeira – Faculdade de Ciências da Vida
16	Universidade de Aveiro – Departamentos de Biologia E Ambiente e Ordenamento
17	Universidade de Évora – Escola de Ciências e Tecnologia
18	Universidade de Trás-os-Montes e Alto Douro – Esc. De Ciências Agrárias e Veterinárias E de Ciências da Vida e do Ambiente
19	Universidade do Algarve – Faculdade de Ciências e Tecnologia
20	Universidade dos Açores – Dep. De Biologia E Ciências Agrárias E Oceanografia e Pescas
21	Instituto Politécnico de Beja – Escola Superior Agrária e Botânico
22	Instituto Politécnico de Bragança – Escola Superior Agrária
23	Instituto Politécnico de Castelo Branco – Escola Superior Agrária
24	Instituto Politécnico de Coimbra – Escola Superior Agrária
25	Instituto Politécnico de Leiria – Escola Superior de Turismo e Tecnologia do Mar
26	Instituto Politécnico de Portalegre – Escola Superior Agrária
27	Instituto Politécnico de Santarém – Escola Superior Agrária
28	Instituto Politécnico de Viana do Castelo – Escola Superior Agrária
29	Instituto Politécnico de Viseu – Escola Superior Agrária
	Scientific Societies and NGOs
30	Assoc. Port. De Estudantes de Biologia
31	Associação Portuguesa de Primatologia
32	Associação Portuguesa para o Estudo e Conservação de Elasmobrânquios
33	Biodiversity4all
34	Carnivora
35	FAPAS
36	Geota
37	Grupo Lobo
38	ICNF
39	Instituto Português de Malacologia
40	LPN
41	Museu da Lourinhã
42	Naturdata
43	Ordem dos Biólogos
44	Projecto Delfim – Centro Português de Estudos de Mamíferos Marinhos
45	Quercus
46	SCIAENA - Associação de Ciências Marinhas e Cooperação
47	Sociedade de História Natural
48	Sociedade Ibérica de Ictiologia
49	Sociedade Portuguesa de Botânica
50	Sociedade Portuguesa de Ciência do Solo
51	Sociedade Portuguesa de Ciências Florestais
52	Sociedade Portuguesa de Entomologia
53	Sociedade Portuguesa de Espeleogia
54	Sociedade Portuguesa de Etologia
55	Sociedade Portuguesa de Fisiologia Vegetal
56	Sociedade Portuguesa de Fitopatologia
57	Sociedade Portuguesa de Herpetologia
58	Sociedade Portuguesa de Microbiologia
59	Sociedade Portuguesa de Vida Selvagem
60	SPEA – Sociedade Portuguesa para o Estudo das Aves
61	SPECO
62	Tagis – Centro de Conservação das Borboletas de Portugal
63	WWF

Annex VIII.
The event on World Wide Web promoted by other institutions/organisations

Instituto Superior de Agronomia's website
(https://www.isa.ulisboa.pt/)*

Fundação para a Ciência e Tecnologia's website
(http://www.fct.pt)*

SPEN's Facebook timeline
(https://www.facebook.com/sociedadeportuguesadeentomologia)[*]

SPEN - Sociedade Portuguesa de Entomologia
15 April · 🌐

The 8th European GBIF Nodes Meeting will take place in Lisbon, between the 19th and 21st of April. The event is organised by GBIF Portugal and will take place at the host institution of the node, Instituto Superior de Agronomia. The meeting is open to all European GBIF participants.

Two side events will take place: a workshop on Sample-based data publication on the 18th of April, and an excursion to the Arrábida Natural Park on the 22nd of April. See more details on the programme page.

http://www.gbif.pt/EuropeanNodesMeeting

> ## Home | gbif.pt
> The 8th European GBIF Nodes Meeting will take place in Lisbon, between the 19th and 21st of April. The event is organised by GBIF Portugal and will take place at the host institution of the node, Instituto Superior de Agronomia. The meeting is open to all European GBIF participants.
>
> GBIF.PT

[*] Sociedade Portuguesa de Entomologia is a Scientific society in the field Entomology

Carnivora's Facebook timeline
(https://www.facebook.com/Carnivora-N%C3%BAcleo-
de-Estudos-de-Carn%C3%ADvoros-e-seus-
Ecossistemas-591656907527299/) *

Carnivora - Núcleo de Estudos de Carnívoros e seus Ecossistemas
24 April ·

O ramo português do GBIF (Sistema Global de Informação sobre a Biodiversidade) organizou na semana passada a 8ª Reunião de Nós Europeus no Instituto Superior de Agronomia, em Lisboa.

Descubra mais sobre este evento e sobre os recursos do GBIF para a comunidade científica e outros potenciais interessados na sua utilização (ex: ONGs).

Home | gbif.pt

The 8th European GBIF Nodes Meeting will take place in Lisbon, between the 19th and 21st of April. The event is organised by GBIF Portugal and will take place at the host institution of the node, Instituto Superior de Agronomia. The meeting...

GBIF.PT

---

* Carnivora is an NGO devoted to the study and dissemination of scientific knowledge regarding Mammalian Carnivores and their ecosystems in Portugal

Annex IX.
Meeting and Workshop Evaluation Forms

Meeting evaluation form

8th European GBIF Nodes Meeting - Evaluation Form

8th European Nodes Meeting

http://www.gbif.pt/EuropeanNodesMeeting/programme

1. General Characteristics - Please rate your satisfaction on the different aspects listed below:

	n/a	1	2	3	4	5
Information sharing	○	○	○	○	○	○
Setting common priorities	○	○	○	○	○	○
Opportunities to interact with fellow peers	○	○	○	○	○	○
Opportunities to voice your personal opinions	○	○	○	○	○	○
Strengthening existing and establishing new collaborations	○	○	○	○	○	○
Information on topics to be discussed and setting of the meeting agenda	○	○	○	○	○	○
Answering of existing or new doubts raised by the discussion	○	○	○	○	○	○
Usefulness for future work at my workplace	○	○	○	○	○	○

(1 = very dissatisfied, 5 = very satisfied)

2. Venue and Organization - Please rate your satisfaction on the different aspects listed below:

	n/a	1	2	3	4	5
Adequacy of the venue	○	○	○	○	○	○
Suitability of the furniture and equipment	○	○	○	○	○	○
Logistics and practical organization	○	○	○	○	○	○
Helpfulness of the organizers before and during the meeting	○	○	○	○	○	○
Meeting participant packet	○	○	○	○	○	○
Lunch and Coffee Breaks (overall)	○	○	○	○	○	○
Hosted Dinner (overall)	○	○	○	○	○	○
Ice-Breaker (overall)	○	○	○	○	○	○
Excursion (overall)	○	○	○	○	○	○

(1 = very dissatisfied, 5 = very satisfied)

3. Objectives - Please assess the attainment of the Meeting objectives listed below:

	n/a	1	2	3	4	5
Defining European Nodes contribution to GBIF's Strategic Plan 2017-2021	○	○	○	○	○	○
Updating from the Secretariat on key subjects	○	○	○	○	○	○
Updating/Reporting by European Nodes collaborators regarding collaborations and new funding opportunities	○	○	○	○	○	○
Shaping European Nodes Work plan 2016-2017	○	○	○	○	○	○

(1 = not attained at all, 5 = fully attained)

4. Best Practice exchange – Thematic break-out group:

Please state what group you were in during Parallel session 1:

○ Nodes portals (and other tools)

○ Sample-based data

How do you assess your group in terms of:

	n/a	1	2	3	4	5
The preparation in advance	○	○	○	○	○	○
The engagement in the discussion	○	○	○	○	○	○
The session facilitation	○	○	○	○	○	○
The depth of the discussion	○	○	○	○	○	○
The outcomes of the discussion	○	○	○	○	○	○

(1=very negatively, 5=very positively)

Please state what group you were in during Parallel session 2:

○ Uses of data, user-science-policy interface and reaching out to the cultural domain
○ Data quality, validation

How do you assess your group in terms of:

	n/a	1	2	3	4	5
The preparation in advance	○	○	○	○	○	○
The engagement in the discussion	○	○	○	○	○	○
The session facilitation	○	○	○	○	○	○
The depth of the discussion	○	○	○	○	○	○
The outcomes of the discussion	○	○	○	○	○	○

(1=very negatively, 5=very positively)

5. How do you assess the discussion around the next action plan topics

Digitization and data repatriation

	n/a	1	2	3	4	5
The preparation in advance	○	○	○	○	○	○
The engagement in the discussion	○	○	○	○	○	○
The session facilitation	○	○	○	○	○	○
The depth of the discussion	○	○	○	○	○	○
The outcomes of the discussion	○	○	○	○	○	○

(1=very negatively, 5=very positively)

Biodiversity informatics curriculum

	n/a	1	2	3	4	5
The preparation in advance	○	○	○	○	○	○
The engagement in the discussion	○	○	○	○	○	○
The session facilitation	○	○	○	○	○	○
The depth of the discussion	○	○	○	○	○	○
The outcomes of the discussion	○	○	○	○	○	○

(1=very negatively, 5=very positively)

6. How do you assess the discussion around the strategic aspects, namely in terms of:

	n/a	1	2	3	4	5
European Nodes Work plan 2016-2017	○	○	○	○	○	○
European Nodes contribution to the launch of a new fitness-for-use task group on Invasive Alien Species	○	○	○	○	○	○
New GBIF Members in the European region	○	○	○	○	○	○

(1=very negatively, 5=very positively)

7. Do you have any other comments and/or suggestions for improving the European Nodes Meeting in the future?

[   ]

OPTIONAL: Please provide your name for any follow ups based on your feedback.

[   ]

[Submit]

Workshop evaluation form

Workshop on Sample-based data publication - Evaluation Form

Workshop on Sample-based data publication – 8th European Nodes Meeting Side Event

http://www.gbif.pt/EuropeanNodesMeeting/workshop

1. Please rate your satisfaction on the workshop aspects listed below.

	n/a	1	2	3	4	5
Content	○	○	○	○	○	○
Structure and schedule	○	○	○	○	○	○
Facilitation	○	○	○	○	○	○
Interaction with my peers	○	○	○	○	○	○
Pre-workshop administration	○	○	○	○	○	○
Adequacy of the venue and its equipment	○	○	○	○	○	○
Practical organisation	○	○	○	○	○	○
Methods/approach	○	○	○	○	○	○
Results and their applicability in my work place	○	○	○	○	○	○
Overall success	○	○	○	○	○	○

(1 = very dissatisfied, 5 = very satisfied)

2. Please rate your level of satisfaction with the workshop stages.

	n/a	1	2	3	4	5
Introduction and introductory presentations	○	○	○	○	○	○
Group work sessions	○	○	○	○	○	○
Group work results presentation and discussion	○	○	○	○	○	○
Workshop wrap up and Conclusion	○	○	○	○	○	○

(1 = very dissatisfied, 5 = very satisfied)

3. What did you find the most useful for your daily work?

4. Have you provided a sample event dataset for the group work session?

○ Yes

○ No

5. If you provided example sample event datasets for the group work sessions did you find it helpful?

6. If you didn't, what was the reason?

☐ No dataset available in your community

☐ Not enough time to prepare a dataset

☐ Instructions how ro prepare the dataset not clear

☐ Other reason (explain)

7. Is there anything that you would have added to this workshop? Any follow-up after it?

8. Do you recommend continuing the practice of organizing a workshop together with the regional European GBIF Nodes meetings? Lessons learn from this workshop or suggestions for topic for a node workshop next year?

9. Any other comments or suggestions?

OPTIONAL: Please provide your name for any follow ups based on your feedback.

Submit