

## MID-TERM ACTIVITY REPORT

*Guidelines on how to complete the activity report are included in italics.*

### Contents

1. Project overview .....	1
2. Executive summary .....	2
3. Introduction to the project and its objectives .....	3
4. Project deliverables - status update by mid-term .....	4
5. Project calendar - status update by mid-term.....	7
6. Project communications .....	8
7. Mid-term evaluation findings and recommendations for the remaining project implementation period .....	8
8. Annex 1 – Additional sources of verification.....	10

### 1. Project overview

*Please insert text in the table*

<b>Project title</b>	Usage of the Atlas of Living Australia (ALA) and GBIF data portals: how to prepare your data, and reuse them on national and international platforms (Workshops)	
<b>Project ID</b>	CESP2018-005	
<b>Project period</b>	<b>Start date (dd/mm/yy):</b> 01/06/2018	<b>End date (dd/mm/yy):</b> 31/05/2019
<b>Project lead</b>		
<b>Name of the main contact person</b>	Anne Bruneau	
<b>Name of institution</b>	Institut de recherche en biologie végétale, Université de Montréal	
<b>GBIF Participant represented</b>	Canadensys	

<b>Role(s) in this project</b>	Mentor / Coordinator	
<b>Project partner(s)</b>		
<b>Name and institution of the representative</b>	<b>GBIF Participant represented (if applicable)</b>	<b>Role(s) in the project</b>
Carole Sinou, Université de Montréal	Canadensys	Trainer / Mentor / Coordinator
Cristina Valverde, Katia Cezón, Miguel Vega	GBIF Spain	Mentor
Jeannette Whitton, Wayne Maddison, UBC	Canadensys	Participants / Co-coordinators
Deborah Paul, Florida State University	iDigBio	Trainer / Mentor
Alain Belliveau, Allison Walker, Acadia University	Canadensys	Participants / Co-coordinators
Mary Kennedy	OBIS	Trainer
David Shorthouse, Canadian Museum of Nature	Canadian Museum of Nature	Mentor / Participants
James Macklin, Agriculture and agrifood Canada	CBIF	Mentor / Participants
David Bloom, VertNet	VertNet	Mentor

## 2. Executive summary

*Provide a brief explanation of the context and the approach taken up to the mid-term of the project, as well as a summary of the main conclusions, lessons learned and recommendations for the remaining project period.*

The main goal of the project is to share knowledge and protocols about good practices in data publication and data-usage, through a series of workshops at different locations in Canada, and to publish documentation about publication and visualization of data on Canadensys and GBIF. In more detail, the aims of the workshops are to teach participants how to publish their data, how to clean and enrich them through a series of tools (Open Refine, APIs, GBIF tools, Georeferencing tools), how to search and download data from the Canadensys (ALA) or GBIF platforms, and what to do with the Biodiversity data available on these platforms.

Canadensys has recently moved to a new framework built upon the Atlas of Living Australia (ALA) framework, enabling new functionalities and tools to index, visualize and analyze data published on the Canadensys IPT, and other Canadian IPTs. This new framework is more powerful than the

previous Explorer available on Canadensys, but is less intuitive, thus leading to the need for user training.

During the first half of the project, we built the material for the workshops and gave two of the planned workshops, one in Nova Scotia (Acadia University, Wolfville) and one in British Columbia (University of British Columbia, Vancouver). The material focuses on data publishing, data-quality enhancement (data-cleaning, data-enrichment, metadata), usage of the ALA Explorer (i.e. Canadensys Explorer) and GBIF Explorer, and data-usage.

The community received positively the idea of holding the proposed workshops, and all participants declare being more skilled after the three-day workshop (see post-workshop surveys listed in section 8). The need for this type of event is evident, but in Canada several potential participants reported that the financial support was insufficient to attend. In terms of the budget, our initial evaluation of the needs was accurate, but the time allocated to the organization of the workshops and preparation of the material was underestimate. The second half of the project will mostly focus on revision of the material, preparation and publication of the documentation, and building of the intended knowledge network. One major point to keep in mind during that second part is to develop our communication channel in order to reach more efficiently our current and potential communities.

---

### 3. Introduction to the project and its objectives

1. *Provide a summary of the overall project to help readers understand its objectives, including, for example:*

- *The project's main objective(s) and expected impact*
- *How the deliverables will contribute to the main objective(s)*
- *The targeted capacity needs as outlined in the project proposal*

2. *Explain to readers what they will find in this report. It should include:*

- *A description of how the evaluation has been carried out (e.g. consultation or surveys with project partners and participants). Please refer to section 3.6 "Monitoring and evaluation" in your original project proposal.*
- *A description of how the project partners will use the evaluation results.*

Canadensys is an active member of GBIF and has published biological data from Canadian institutions (universities, museums, ministries, non-profit organisations, citizen-science projects) since 2011. In 2016, Canadensys tested and installed basic modules of the Atlas of Living Australia (ALA) framework to replace its Explorer, which had reached its technological limits. Now that the basic ALA modules (Explorer, Collectory, Images) are installed and running, we need to train users to fully take advantage of this new framework.

In addition, the GBIF data portal, updated in 2017, offers a large number of tools and features to facilitate cleaning, curation, discovery and analysis of biodiversity data. However, in discussions with users, we realized that a large number of these tools and features are not known or used by researchers, curators and students who would benefit from knowing more about them.

In order to maximize the quality of data published on both the GBIF and Canadensys platforms, and to encourage use of data for scientific research, we need to offer opportunities for our community of users and partners to learn and experience first-hand how an ALA framework and GBIF data portals work.

In coordination with GBIF Spain and iDigBio, we planned a series of workshops in Canada to train users on how to best take advantage of all the features offered by these portals. The material produced for the workshops and step-by-step guides will be made freely available in English, French and Spanish, and will be reusable by other countries and institutions. All the material and documents produced during the project will serve as good-practices documentation, giving tools and protocols to users of IPT, ALA framework and GBIF platform.

The evaluation of the first half of the project was done through surveys sent to participants of the workshop, both before and after the event. The evaluation of the data quality enhancement is not possible at this stage of the project, since participants have not done updates of the current datasets or published new ones.

The surveys evaluation will help us improve the material for the remaining workshop and before publishing it on other platforms. It will also help us define our future projects since one of the questions was about the needs and dreams of participants regarding biodiversity data publication and usage.

---

#### 4. Project deliverables - status update by mid-term

*Provide updates on the status of each of the planned deliverables by the mid-term, listing the related activities connected to these and explanatory notes. In the event of identifying a deliverable as "delayed", please provide detailed explanatory notes and planned actions to address this issue. Please provide links to any documents or webpages in the "Sources of verification" column. Add as many rows as needed.*

Deliverable	Related activity	Status: Completed / on track / delayed	Explanatory notes	Sources of verification
<p>Workshop documentation and other materials (theory and practical exercises) about:</p> <ul style="list-style-type: none"> <li>• ALA framework (based on the Canadensys portal) and IPT <ul style="list-style-type: none"> <li>○ Step-by-step: how to publish data (Darwin Core, IPT, metadata)</li> <li>○ How to use the new Canadensys Explorer (or other ALA portal).</li> </ul> </li> <li>• GBIF Portal <ul style="list-style-type: none"> <li>○ For users: Search for specific taxa, or regions and download data. Use of the GBIF backbone.</li> <li>○ For Data Managers: usage of GBIF metrics and Data Validation Tools to assess data quality.</li> </ul> </li> <li>• Tools <ul style="list-style-type: none"> <li>○ OpenRefine (API)</li> <li>○ Data Cleaning with R</li> <li>○ GBIF Species matching</li> </ul> </li> </ul>	<p>Workshops in Wolfville (Acadia University, 14-16 October 2018), Vancouver (University of British Columbia, 7-9 November 2018) and Ottawa (date to be determined)</p>	<p>Completed (but need revisions)</p>	<p>The material was created for the workshops. A revision is necessary prior to online publication.</p>	<p><a href="https://drive.google.com/open?id=16XY7wP0xzK07fQ2vGpQAYxwa3bRMsOUm">https://drive.google.com/open?id=16XY7wP0xzK07fQ2vGpQAYxwa3bRMsOUm</a></p>
<p>Translation of material in French and Spanish, possibly Portuguese</p>	<p>Workshops</p>	<p>On track</p>	<p>Planned for the second half of the project</p>	
<p>Publication of the presentation (PDF) and recording of the workshops (on Canadensys community page, Living Atlases</p>	<p>Workshops</p>	<p>On track</p>	<p>Planned for the second half of the project. The workshops in Nova Scotia and</p>	

Deliverable	Related activity	Status: Completed / on track / delayed	Explanatory notes	Sources of verification
community page, partner networks, i.e. GBIF Spain, iDigBio, VertNet, etc.)			Vancouver were not recorded.	
Publication of a 'mentoring-map' in the Canadensys network, with contact information of the new mentors trained (participants at the workshops), thus enhancing the spread of knowledge within the community	Workshops	On track	Planned for the second half of the project. In the post-workshop survey, we asked participants if they wished to be part of the knowledge network.	<a href="https://docs.google.com/spreadsheets/d/1u9dKP913_MLFB_eZhEh3wrHoNj1myzHhp6K_PlzAgNFk/edit?usp=sharing">https://docs.google.com/spreadsheets/d/1u9dKP913_MLFB_eZhEh3wrHoNj1myzHhp6K_PlzAgNFk/edit?usp=sharing</a>
Development or update of step-by-step PDF guides, complementary to the videos developed during the CESP2018-004 project, in French, English and Spanish, about: <ul style="list-style-type: none"> <li>• How to search the ALA portals.</li> <li>• How to clean data. What are the best tools and ways to do this.</li> <li>• How to use OpenRefine to prepare and clean data.</li> <li>• Revision of the Canadensys <a href="#">7-step-guide to data publication</a>.</li> </ul>	Workshops	On track	Planned for the second half of the project	
Publication of new datasets on Canadensys and GBIF	Workshops	On track	Publication of new datasets will probably start in the second half of the project.	

## 5. Project calendar - status update by mid-term

The calendar should be completed the same way as in section 3.7 of your Full Project Proposal, but should also clearly indicate any expected changes (e.g. use of colour-coding to indicate expected changes). Please provide justifications for any expected changes in the notes column.

Activity	2018							2019					Notes
	J	J	A	S	O	N	D	J	F	M	A	M	
Preparation of the workshop material	X	X	X										
Planning of the workshops	X	X	X	X	X	X							
8-10 August 2018: Workshop #1 (Ottawa, ON)			X							X	X		Postponed in late winter / early spring, due to time constraints
24-26 September 2018: Workshop #2 (Wolfville, NS)				X	X								Change of date to accommodate other on-going projects
7-9 November 2018: Workshop #3 (Vancouver, BC)						X							
Revision of the workshop material			X		X		X						
Deadline for mid-term reporting						X	X						
Writing of documentation								X	X				
Translation of material and documentation in French and Spanish									X	X			
Publication of material and documentation online										X			
Final evaluation survey + analyses of the answers											X		
Deadline for final reporting												X	

---

## 6. Project communications

*Describe how the results of your project are being, and will be shared with the project stakeholders and broader GBIF community.*

*Please also review the project webpage available from <https://www.gbif.org/programme/82219/capacity-enhancement-support-programme#projects> and provide any additional documents, events, news items or links that you would like to add to your page.*

The workshops was advertised on the Canadensys Community website

(<http://community.canadensys.net/2018/how-to-prepare-your-data-for-publication-and-what-to-do-with-open-data-workshop>, <http://community.canadensys.net/2018/how-to-prepare-your-data-for-publication-and-what-to-do-with-open-data-vancouver-workshop>), and on our Twitter and

Facebook feeds. The workshop material will be made available from these pages, as well as from the project page in the CESP pages. Twitter coverage during the Vancouver workshop was vivid, generating interesting discussion about recognition of the important role of data provider and the need for recognition of this role (<https://twitter.com/idbdeb/status/1060230807736860672>)

Further communications linked to this project will be through the same path (blog, social media), as well as other platforms (Living Atlases Community documentation).

Following feedback from participants, we also decided to start a regular publication of “how to” blog posts, with simple recipes for data cleaning with Open Refine, data enhancing through different tools (API, GBIF tools like Species Matching, georeferencing tools).

---

## 7. Mid-term evaluation findings and recommendations for the remaining project implementation period

*This section is to summarize and document the mid-term evaluation of the project's progress and any changes to the project plans based on the findings of this mid-term evaluation.*

*This section of the report should also cover, for example:*

- *Areas of success/best practices/lessons learnt from this first stage of project implementation, which can be built on during the remainder of the project period.*
- *Any feedback on the project's relevance from the partners and stakeholders.*
- *Comments on the project implementation, its efficiency and effectiveness.*
- *The management arrangements for the project, including support from the GBIF Secretariat.*
- *Any reflection on the mid-term evaluation itself that could help inform the project's final evaluation and final report.*

*Please also use this section as an opportunity to identify the main lessons from the project experience that could be applied in other contexts, including any best practice that others in the GBIF community could apply. In addition, outline any recommendations for the GBIF Secretariat or the community to reinforce the initial success of the project.*



The CESP project as proposed is going well, is on track, and the projected objectives have been attained. Two of the three workshops have taken place, and the third one, in Ottawa, has been postponed to the second half of the project, due to time constraints. The workshop material is available for the workshop participants and will be made available to everyone after revisions, during the second term of the project.

We are highly grateful for the help received from Mary Kennedy (OBIS Canada) and Deborah Paul (iDigBio) during the workshops in Wolfville and Vancouver. We also thank the GBIF Secretariat for sharing course material on data-quality and data publication. These courses were used as baseline documentation for the workshop material for this project.

A total of 6 participants attended the workshop in Wolfville (Acadia University). The aim to attract 20 participants was unfortunately not attained, but this smaller group allowed for more direct support for each participant. The Maritimes region is vast, and the cost to attend the workshop appears to be a drawback for several people. Furthermore, some people demonstrated interest in the workshop but were unable to attend due to academic duties (teaching, meetings, deadlines). This problem highlights the need for access to relevant documentations on Canadensys and other platforms. Another explanation for the low attendance during this workshop is explained by the fact that several people from Gaspésie (Québec) demonstrated their interest for this workshop but would prefer to attend a workshop in French. In order to answer to this need, a fourth workshop will be organized in French in Mont-Joli (Québec) in 2019.

In Vancouver, a larger group attended the workshop (18 people), with participants from British Columbia, Alberta, Washington State (US) and Florida (US). This workshop was held as an unofficial satellite meeting of the Entomological Society of American Annual Conference.

For both events, participant' feedback is positive and everyone agrees that this type of event is relevant in order to give access to more standardized biodiversity data. However, a participant from the Vancouver workshop noted in the post-workshop survey that we should spend more time on data-usage. Due to time constraints, we spent less time on that topic, but this comment shows the interest for another workshop on data-usage, or for the addition of material for one more day on data-usage.

The workshop in Vancouver raised the possibility of organizing other events near the Canada-USA boarder, strengthening the partnership between iDigBio and Canadensys.

Overall, the mid-term objectives of the project have been reached and the community reacted favourably to the project. This workshop project has allowed the initiation of discussions about new partnerships and potential other projects.

## 8. Annex 1 – Additional sources of verification

*Please use this section to provide any additional sources of verification, for example links to relevant digital documents, news/newsletters, brochures, copies of agreements with data holding institutions, workshop related documents, pictures, etc.*

### Workshop material:

- Wolfville:  
[https://drive.google.com/open?id=17wfuonwz\\_4Mp4aGWTGcMTAhJ8LzUGfUp](https://drive.google.com/open?id=17wfuonwz_4Mp4aGWTGcMTAhJ8LzUGfUp)
- Vancouver:  
<https://drive.google.com/open?id=16XY7wP0xzK07fQ2vGpQAYxwa3bRMsOUm>

### Pre and Post-workshop surveys summary:

- Wolfville:  
[https://docs.google.com/spreadsheets/d/1u9dKP913\\_MLFBeZhEh3wrHoNj1myzHhp6K\\_PlzAgNFk/edit?usp=sharing](https://docs.google.com/spreadsheets/d/1u9dKP913_MLFBeZhEh3wrHoNj1myzHhp6K_PlzAgNFk/edit?usp=sharing)
- Vancouver:  
[https://docs.google.com/spreadsheets/d/1Eeu80BWZrOBS1RBIISiThrAyg216tRIPiNHg\\_hwlxNuU/edit?usp=sharing](https://docs.google.com/spreadsheets/d/1Eeu80BWZrOBS1RBIISiThrAyg216tRIPiNHg_hwlxNuU/edit?usp=sharing)

### Twitter:

- Workshop in Wolfville: <https://twitter.com/Canadensys/status/1051942542252666881>
- Workshop in Vancouver:
  - o <https://twitter.com/Canadensys/status/1060647107017695232>
  - o <https://twitter.com/jessilynnallen/status/1061133336360321024>
  - o <https://twitter.com/JohannaJantzen/status/1061059490299158533>
  - o <https://twitter.com/idbdeb/status/1060215800068366336>