Annual Report
IAVS Working Group for Ecoinformatics
July 2018

Stephan Hennekens (chair), Florian Jansen, Peter Minchin & Susan Wiser

The goals of the Ecoinformatics Working group are to facilitate communication among scientists studying community ecology through exploration and synthesis of large databases comprising vegetation-plot and related ecological data; facilitate access to these data; establish standards for exchange of these data to facilitate data sharing and to provide tools for identification, access, integration, storage, and analysis of these data. To meet the requirements to be a recognized IAVS Working Group and be eligible for IAVS funding, a set of Bylaws were drafted in February 2013, based on those of the European Vegetation Survey Working Group, were approved by members of the working group in March 2013 and approved by the IAVS Council at the Tartu symposium in June 2014.

Concerning the Ecoinformatics website only minor changes have been made last year. Parts of the former website (hosted at the University of North Carolina) have been moved to the IAVS website, but there is much more information that can be provided to the Ecoinformatics community. A proposal for the update was submitted by the Steering Committee in 2016 last year, but rejected by the IAVS Governing Board. Last year the lead person of the proposal, Florian Jansen, has been too busy to revise the proposal. Nevertheless it would be good to put the website revision on the agenda again for next year.

The meeting of the Ecoinformatics Working group in Palermo last year was very inspiring and broad up quite a few ideas, of which at least one has been successfully worked out (development of R-tools for Veg-X). The minutes are included in the annex.

In 2003, the Working group decided to promote the development of a standard to exchange vegetation-plot data. An international collaboration produced a draft standard, Veg-X for data exchange, implemented in an XML Schema (Extensible Markup Language). Veg-X was designed to be compatible with the most commonly used vegetation databases.
In line of this initiative a proposal, that was granted and successfully completed, was the development of a R-package to migrate vegetation plot data into the Veg-X exchange standard format. Miquel De Cáceres, Sebastian Schmidtlein and Susan Wiser therefore received a small grant last year from IAVS to develop an R-package for migrating vegetation plot data to Veg-X compatible format. An article on this will be published in one the IAVS Bulletins soon. The files conforming the Veg-X XML schema (version 2.0) and the R package *VegX* can be downloaded from a GitHub repository (https://github.com/miquelcaceres/VegX). A detailed description of the schema and a user manual of the package can be found at https://miquelcaceres.github.io/VegX/.
Worthwhile to mention is that the very first tool for collecting field data with Veg-X as export format is the Android app Vegapp, that is about to be published on Google Play. Such an application as Vegapp facilitates the transformation of data in the Veg-X standard for users who are unfamiliar with R. A widespread adoption of Veg-X and related import and

export tools will contribute substantially to the exchange, harmonization and reuse of vegetation plot data.

Last year one of the Working Group members, Miguel Alvarez, has developed a package for retrieving and handling vegetation-plot databases in R, and is published on CRAN (Comprehensive R Archive Network). More details on this package and the dependency 'taxlist' are on the respective sites:
https://github.com/kamapu/vegtable
https://github.com/kamapu/taxlist

After four years of being active the Steering Committee needs to be replaced by a fresh new team. Therefore an election among all members of the WG is set up.

Annex: Ecoinformatics WG meeting 2017-06-23 in Palermo

All three working group goals mentioned in the WG Bylaws are still fine:
(1) to establish standards for data exchange to facilitate data sharing,
(2) to provide tools for data identification, access, integration, storage, and analysis, and
(3) to facilitate communication among scientists studying community ecology through exploration of multiple large databases.

In all three areas there have been substantial achievements in the last decade, more or less connected to the WG (although not necessarily initiated by the WG).

Several ideas arose during the WG meeting about activities that could be undertaken by the working group.
- organise the election of a new WG steering committee before the end of 2017
Explanation: Who should conduct this to be somewhat independent of current steering committee? Given that we need to plan activities for the next year, should the transition to a new steering committee take place at the IAVS meeting in Bozeman? We need to decide on the timing to a) solicit nominations from the WG members; and b) carry out voting. Sebastian Schmidtlein coordinated our last election. There also may be good procedures established by other WG (e.g. EDGG).

- develop a more informative webpage
Explanation [see associated proposal]

- help to establish and advance (sub)continental mega-databases (TAVA, LAVA/SAVA, Asia etc.)
Explanation This was suggested by Juergen Dengler. Perhaps he can take the initiative to describe this more completely and outline what the next steps would be.

- to develop 'good practice' guidelines for creating smaller databases.
Explanation Do we envision some kind of document describing the workflow? Are there any examples we can follow? This would link to the template listed next

- develop a template for vegetation plot data
Explanation: Data capture templates could be as simple as well-structured spreadsheets or database UI (user interfaces) to capture the primary types of vegetation plot data (e.g. site data, releves, frequency transects, count data measures of individuals (e.g. forest inventory data), repeated measures of both. This would also require some sort of 'back end' system for storage or the ability to easily contribute these data to a suitable storage located (e.g. VegBank). We need to consider solutions for the entire workflow. [There is a link to ideas expressed regarding Veg-X.]

- organise an IAVS conference workshop for vegetation data preparation workflow (e.g. in parallel to the planned Biome workshop)

Explanation: This would be oriented to meeting the needs of people who are new to this aspect of vegetation science, especially Young Scientists. We will need to interact with the Young Scientists working group to formulate the content of this workshop. Who is willing to lead this? Probably the 'good practice' guidelines described above need to be completed (at least in a draft form) first.

- write guidelines for data storage possibilities if this is requested for publications

Explanation: Veg-X could be promoted as standard for storing data underpinning publications in IAVS journals. Tools for transforming user-defined formats to Veg-X (e.g. through an R package) should be provided and tested before encouraging authors to archive their data in this format. (see Veg-X section)

- secure a backup of vegetation databases (IAVS data save)
Explanation: This was suggested by Bob Peet as a fourth goal for the working group. Currently all of the large-scale databases and synthesis efforts (e.g. BIEN, sPlot) do not have long-term secure funding. What happens if they lose their funding? Is it possible for IAVS to develop a means for archiving these data? Or can IAVS investigate a suitable cloud-based archive solution? Could Bob be an appropriate person to lead such an investigation?

- use the special ecoinformatics section in Phytocoenologia to publish articles
Explanation: Long or Short (1-page) Reports about vegetation databases are already well established and can be found in nearly every issue of Phytocoenologia. But also other topics like standards discussed by the working group or methodological papers can be submitted.

- work on tools for 1) data entry 2) data exchange 3) data storage
Explanation: Are 1) and 3) covered already? Also see Veg-X section. @Florian: I think we have preliminary solutions for all three areas but depending on experience or region they might be difficult to apply and more tools for all three areas would be useful.

- Data exchange -- advancing Veg-X
Explanation: Susan, Sebastian and Miquel met after the meeting and discussed this. As the Veg-X schema and documentation is no longer supported on the TDWG website, it would be good to provide this on the IAVS WG website.
As a first step we need to determine which version of Veg-X should be the 'official version'. Sebastian's app is written against 1.5.1. There is documentation of the changes between 1.5.1 and 1.5.2. There is a version 1.5.3 that appears to have incorporated suggestions from the BIEN WG. Susan is attempting to sort out whether there is documentation supporting 1.5.3 so we can decide on the 'official' version to post as this needs to be stable.

Then a workflow utilising Veg-X needs to be developed to support the following workflow:
(1) Data collection - Vegetation Survey App (e.g. what Sebastian has developed) - needs extension to individuals and repeated measures. It would then export to Veg-X.
(2) R package that allows writing/reading XML of Veg-X. It has R structure [in memory] that replicates Veg-X. From this structure you could also extract data in particular formats (individual-based, stratum-based, and relevé). Miquel expressed interest in developing such an R package. This package would help authors to transform their field data, originally collected and stored in various data structures, into a format appropriate for archiving and exchanging.
(3) Extracts could include an XML document that can be archived in a suitable archive.
Suggest that people archive data in format when publishing through IAVS journals. More structured databases (TurboVeg3, VegBank) would need to write imports of this format to enable more ready access to these data via friendlier User Interfaces.