DOI - 10.21570/BUL-201705-2

IAVS and Data for Vegetation Science: Past, Present and Future

Based on a presentation delivered at the 2016 Annual Meeting of IAVS upon being awarded Honorary Membership in the Association

Robert K. Peet

University of North Carolina, Chapel Hill, NC 27599-3280 USA (email: peet@unc.edu)

Introduction

Much of science started out as a search for unifying principles and grand generalizations. Vegetation Science, in contrast, started out as an effort to document and understand the natural variation and diversity in the assemblage of plant communities. In this early focus on natural diversity rather than general principles, it paralleled systematics. Part of the reason for this detail-oriented approach can be found in a quote from Robert May (1986), summarizing the perspective of Robert MacArthur, "…ecology is a science of contingent generalizations, where future trends depend (much more than in the physical sciences) on past history and on the environmental and biological setting". In short, before we can understand the general patterns and processes, we need to have a clear perception of the diversity of responses in space and time so that we can place the details within the context of the critical contingencies.

Vegetation science has come a long way. Description still plays a major role in vegetation science, from local vegetation to global patterns and how they are changing. These data and their information on context of ecological patterns and process helps us greatly in our efforts to explain and understand vegetation. We see many components of this in our journals, including such topics as community assembly and its linkage to traits, biotic interactions, spatial and temporal dynamics and ecosystem processes. This emerging deeper understanding can then be applied to conservation, policy, management and various forms of prediction.

Many vegetation science activities depend on accumulated data, and the dependency has grown ever stronger. As Keeling et al. (2009) suggest, "…this is true because of the complexity of ecological systems, particularly when viewed at large spatial and temporal scales. Data-intensive science organizes large volumes of data from multiple sources and fields…" Other fields with a need for large quantities of data developed informatics resources well before ecoinformatics emerged as a sub discipline of ecology and vegetation science. One reason for this is the broad range of formats, sampling protocols and data structures within the vegetation science community, much of which results from cultural and regional traditions (Jones et al. 2006).

I am particularly sensitive to the role of data in vegetation science as its emergence and development has paralleled my own career, and I have contributed, albeit modestly, in a number of different ways. Here, strongly influenced by my own history of association with IAVS, I describe some of the evolution of data analysis and data resources in the development of modern vegetation science. I then consider the resources and infrastructure likely to be needed as we move forward. Finally, I look at the challenges these present and how IAVS and other professional organizations could and should play a major role in supporting the information infrastructure needed to advance our field in coming decades.

The Emergence of Data Analysis in Vegetation Science

Early years of computing in Vegetation Science

The 1960s saw the first wide-spread availability of computers for applications in ecology. No longer were vegetation scientists constrained to analyze community data on paper or desktop adding machines. Large-scale classification and ordination analyses had become possible. The opportunities were quickly recognized by a subset of IAVS members, but there was little consensus on how best to exploit them and the approaches were numerous.

In 1968 several members of IAVS, largely under the leadership of Sandro Pignatti, proposed formation of a Working Group for Data-Processing in Phytosociology to share methods and facilitate the advance of computer applications in vegetation science (van der Maarel 1971). A series of meetings and symposia followed (van der Maarel 1974). One particularly important change for IAVS was the reorganization of the Association journal Vegetatio in 1974 with a new editorial board that was much more receptive to quantitative analyses and methodological articles. This period culminated in 1979 with a widely-attended meeting of the Working Group hosted by Eddy van der Maarel in Nijmegen (van der Maarel 1980; Fig.1). This was the first IAVS meeting I attended, and, in my perception, the meeting dramatically and permanently changed IAVS from its historical focus on classical phytosociology to a modern professional association with members spanning a broad range of interests and approaches.

Although computer programs were created and shared by ecologists in the 1960s, it was not until the 1970s that vegetation scientists started to converge on standard methods, comparing their characteristics, and creating widely available packages of programs. Particularly influential in this movement was software included in the Cornell packages (e.g. Gauch 1973, Hill 1979). Meanwhile, synthetic works started to appear that compared methods and attempted to guide users (e.g. Orlóci 1978, Gauch 1982, Pielou 1984).

Transition to greater openness, large-scale synthesis, and big data

The growing breadth of interests of IAVS members and their need to communicate among themselves led to another important transition within the Association. The journal *Vegetatio*, although the official journal of the Association, was not owned by the Association. This resulted in lack of control, as well as prices beyond the means of most members and many libraries. In 1989 Eddy van der Maarel, in collaboration with Robert Neuhäusl and myself, created a new journal, owned by IAVS, with modest prices and with most of the same editors that had previously been associated with *Vegetatio*. As Eddy wrote (van der Maarel 1990), this was intended to be “a journal for all vegetation scientists.” This step also demonstrated a growing interest in IAVS with respect to providing critical infrastructure and making information more open.

Increased computational resources for vegetation scientists quickly led to extensive digitization of data and an expectation of larger-scale synthetic works. In particular, national-level syntheses of vegetation were undertaken in a number of European countries, such as Britain (Rodwell 1991-2000), Austria (Mucina et al. 1993) and The Netherlands (Schaminée 1995–99). The combination of this growing number of national-level syntheses and the fall of the Iron Curtain opened new opportunities and inspired work toward a European synthesis (Mucina et al. 1993, Rodwell et al. 1995) following common standards (Mucina et al. 2000). Meanwhile, similar enthusiasm for large-scale synthesis was building in the US. In 1993 several of us convened a symposium held during the annual meeting of the Ecological Society of America (ESA), which catalyzed a national effort to unify the diverse classification initiatives being conducted across the country. This quickly led to

Fig. 1. Symposium on Advances in Vegetation Science, May 1979, Nijmegen, NL. Among the many: M.P. Austin, R.S. Clymo, M.B. Dale, J.B. Faliński, E. Feoli, L.F.N. Fresco, D.C. Glenn-Lewin, G. Grabherr, P. Greig-Smith, V. Komárková, H. Leith, R. del Moral, L. Orlóci, R.K. Peet, S. Pignatti, I.C. Prentice, F. Romane, H. Sjörs, M. Werger, V. Westhoff, O. Wildi & E. van der Maarel.

establishment of the ESA Vegetation Classification Panel. The near simultaneous creation of the ESA Vegetation Classification Panel, the emergence of a US Federal Geographic Data Committee mandate for a national vegetation classification standard, and an initiative of the Nature Conservancy to compile a list of the vegetation types of the country, led to a three-way collaboration starting in 1994 and culminating in the first national compilation of vegetation types (Anderson et al. 1998; see Barbour et al. 2000).

Large-scale national and international analyses require databases for management of the many thousands of vegetation plots involved. To support the national classification initiative of The Netherlands, Stephan Hennekens developed the program TurboVeg (Hennekens 1995, 2001), which was soon employed broadly across Europe. The initial design was kept relatively simple so as to facilitate adoption by multiple groups. In addition, data were assumed to conform to the standard relevé methods used widely in Europe. The program was run on local computers, each supporting its own unique database. Subsequent versions of this program form the backbone for a number of more recent, large-scale initiatives (e.g. Chytrý et al. 2016, Mucina et al. 2016, Walker et al. 2016).

In the US, vegetation ecologists tend to be idiosyncratic in their data collection methods, so adoption of a simple data model based on the relevé method or an equivalent methodology was not a viable option. To solve the problem, the ESA Vegetation Classification Panel asked me to lead an effort to provide a national vegetation data framework. This led to creation of VegBank, first released for public use in 2003 (Duke 2006, Peet et al. 2012). This is a large-scale database, built with open-source software, publically accessible over the web, and which has considerable flexibility in the format of data submitted and served. It is designed to allow users to easily submit, search, view, annotate, aggregate, cite, and download diverse types of vegetation plot data. Unique digital identifiers are assigned to individual plots as well as user-created plot datasets.

Of particular importance to vegetation scientists, and ecologists more broadly, has been the emergence of the new field of ecoinformatics. This was made possible by the growing availability of large datasets, particularly the opportunity to intersect those that contain information about species occurrences, species co-occurrences, and the many types of site data ranging from climate to soils to remotely-sensed information. A necessary first step was to make data more available, which was greatly enhanced by various initiatives for long-term data sharing. For example, as part of an ESA initiative toward open data, I implemented, as Editor-in-Chief of *Ecology* and *Ecological Monographs*, such innovations as digital supplements to journal papers and data papers dedicated to describing and archiving important ecological datasets (Peet 1998). At roughly the same time, a large ecoinformatics initiative was built at the US National Center for Ecological Analysis and Synthesis (NCEAS) in Santa Barbara, California, and its staff have subsequently devoted considerable effort to the development of tools for archiving, interpreting and integrating ecologically relevant data (see Michener & Jones 2012). These, and many parallel initiatives, were finally making it possible to analyze ecological data in such a way as to address the need for contingent generalizations as described by May (1986).

The initiatives at NCEAS, including development of VegBank by our international group of vegetation scientists, led to significantly increased awareness of the importance of ecoinformatics to the vegetation science community. This, in turn, led several of us to organize a full-day session on ecoinformatics at the 2003 IAVS meeting in Napoli. The participants met there and organized a new IAVS Working Group for Ecoinformatics, which was immediately approved by the IAVS Council. Finally, the participants developed a charge to the new Working Group to focus their activities, all components of which are still active concerns of the ecoinformatics community within vegetation science. The charge contained the following components: 1) develop for plot data an international data exchange standard including an XML schema; 2) recommend standards and requirements for archiving plot data; 3) communicate with … other organizations regarding taxonomic database needs; and 4) address issues related to requirements for extended queries, intellectual property rights, and confidentiality.

Data in Contemporary Vegetation Science

Vegetation survey and synthesis

Increased availability of vegetation plot data has greatly accelerated efforts to compile and homogenize large-scale regional classifications. In Europe, the IAVS European Vegetation Survey meets annually. Their efforts have produced a database of well over one million plots across 57 countries (EVA, see Chytrý et al. 2016), plus a synthetic treatment of 1,108 alliances of vascular plant communities spanning all of Europe and adjacent islands (Mucina et al. 2016). In the US a collaboration by ESA and NatureServe led to the Federal Geographic Data Committee adopting new standards for vegetation classification and coordinated revisions based on plot data (FGDC 2008, Jennings et al. 2009; Fig. 2). The content of the current, 8-level classification has been peer reviewed and an editorial process is in place

Fig. 2. Meeting of the Cyberinfrastructure Design Committee for the US National Vegetation Classification, October 2015. Participants included, left to right, Michael Lee (VegBank Manager, University of North Carolina), Christina Justice (Consultant, Innovative System Solutions Corporation), Robert Peet (University of North Carolina; Principal Investigator, VegBank Project; Executive Committee, ESA Vegetation Classification Panel), Kristin Snow (Ecology Database Analyst, NatureServe), Cliff Duke (Director of Science Programs, Ecological Society of America), Alexa McKerrow (US Geological Survey; Coordinator, US National Vegetation Classification), and remotely on screen Marianne Burke (US Forest Service; Chair, US Federal Geographic Data Committee Vegetation Subcommittee) and Don Faber-Langendoen (NatureServe; Chair, US National Vegetation Classification Peer Review Board).

for proposed revisions. Currently there are 1,263 alliances and 6,168 associations recognized and described for natural vegetation of the continental US (Faber-Langendoen 2017). Similarly, a large-scale effort is underway to compile plot data across all of the arctic (Walker et al. 2016) and to develop a coordinated classification based on Braun-Blanquet methods (Walker et al. 2017).

Large database initiatives

Many sources of data are now widely available to vegetation scientists including not only plot data, but also environmental data, species distribution data, species trait data, and phylogenetic data. Inevitably, a number of initiatives have developed to integrate and simultaneously analyze these diverse types of data. Two projects I am particularly familiar with are BIEN (Botanical Information and Ecology Network; Enquist et al. 2016) and sPlot (Dengler et al. 2014, Purschke 2015). In the BIEN project we have attempted to bring into one database all open-source species occurrence records, vegetation plot data, trait data and phylogenetic data for plant species of the New World. The data are publicly available, as are derived distribution maps and models for ~92,000 New World plant species. With the sPlot initiative, we are trying to bring together vegetation plot records from across the globe and integrate them with trait data and phylogenetic data. As of 2016 there were over 1.1 million plots in the sPlot database. Both of these projects have multiple working groups generating diverse and innovative synthetic publications. Doubtless there will be other groups of this sort that integrate and analyze the accumulating digital information on plants and vegetation. However, it would be most efficient for vegetation science to maintain and grow a few specific projects to assure long-term preservation and maintenance and to avoid unnecessary costs and redundant efforts.

Future Directions and Opportunities

Addressing challenges associated with big data

The development of large databases for vegetation science not only has provided invaluable resources, but has also made clear a number of challenges that our community should address (Wiser 2016). In many cases solutions have been proposed and at least partially implemented. IAVS could greatly benefit vegetation science, and ecology more broadly, by taking a leadership role in providing solutions to these problems.

1. Data discovery and maintenance
Even if data are digitized, this does not mean they are discoverable or that they will persist. Indices of databases can be very valuable tools. One such tool is GIVD, the Global Index of Vegetation Databases (Dengler et al. 2011). Registration of plot databases in GIVD is voluntary, but already there are 244 databases registered representing in excess of 3.1 million plots. We should encourage the development of more such registries and attempt to assure their long-term maintenance. IAVS could also serve as a home for databases in much the same role as, for example, GenBank serves the genomic community. Just as importantly, many of our current databases lack a clear mechanism for long-term support (e.g. BIEN, sPlot, VegBank). IAVS could work with the managers of such databases in an effort to ensure their continued survival and improvement as their initial, short-term funding fades away, and ultimately to bring their critical data together into a large, global resource.

2. Data formats and protocols
Ecologists across the globe are remarkably inconsistent in their data formats and protocols. Increased consistency in data collection protocols and formatting would be immensely constructive. IAVS should suggest a set of standard protocols and formats. Moreover, we ought to develop consistent vocabularies. The various efforts to develop ecological ontologies have the potential to greatly help in this regard (Madin et al. 2008), as do large database initiatives with carefully assembled constrained vocabularies.

An important initiative undertaken by the IAVS Ecoinformatics Working Group is to develop and maintain data exchange standards so that computer programs and databases with different formats can seamlessly exchange data. A first attempt was VegX, an XML standard for exchange of vegetation plot data in different formats (Wiser et al. 2011). Subsequently, a TDWG Observations Task Group was formed to update VegX, and more broadly to create a core semantic model for observations across the ecological sciences (http://www.tdwg.org/activities/osr/obs/). Completion of this task has the potential to greatly benefit vegetation science, and ecology more broadly.

Although various large groups within IAVS, and vegetation science more broadly, are striving to develop regional synthetic classifications of vegetation, these groups are not coordinated in a way that would facilitate the eventual development of a global synthesis. De Cáceres et al. 2015 recognized this problem and provided a framework for classification activities to facilitate such an eventual global synthesis. The IAVS Vegetation Classification Working Group is discussing how to build on the De Cáceres et al. framework to develop a unified approach for a global vegetation classification. IAVS should support such efforts and encourage greater international coordination and cooperation in vegetation classification.

3. Quality control and consistency
Database content is notoriously prone to errors. For example, BIEN found many thousands of terrestrial plant occurrences with coordinates that map to oceans. Moreover, not only are geo-coordinates often incorrect, species and location names are also prone to inconsistences and errors. Tools are needed to identify and correct such errors in location and taxonomy. Some tools are already available such as BioGeomancer for geographic validation (e.g. Guralnick et al. 2006) and the Taxonomic Name Resolution Service (Boyle et al. 2013) for matching plant names against various authoritative lists. Moreover, the sPlot group has created a separate database for determining whether species are native to a particular region (GloNAF, see Van Kleunen et al. 2015). Still, these represent first steps and we need more powerful tools.

A particularly difficult challenge is presented by situations where there is a many-to-many relationship between names and categories. This is best known in biological nomenclature where one name can refer to many different sets of organisms and one organism can have many different names. This challenge, as well as potential solutions, was first clearly articulated by Berendsohn (1995). Efforts to merge plot data from multiple investigators working in different areas at different times and following different taxonomic authorities have reinforced this

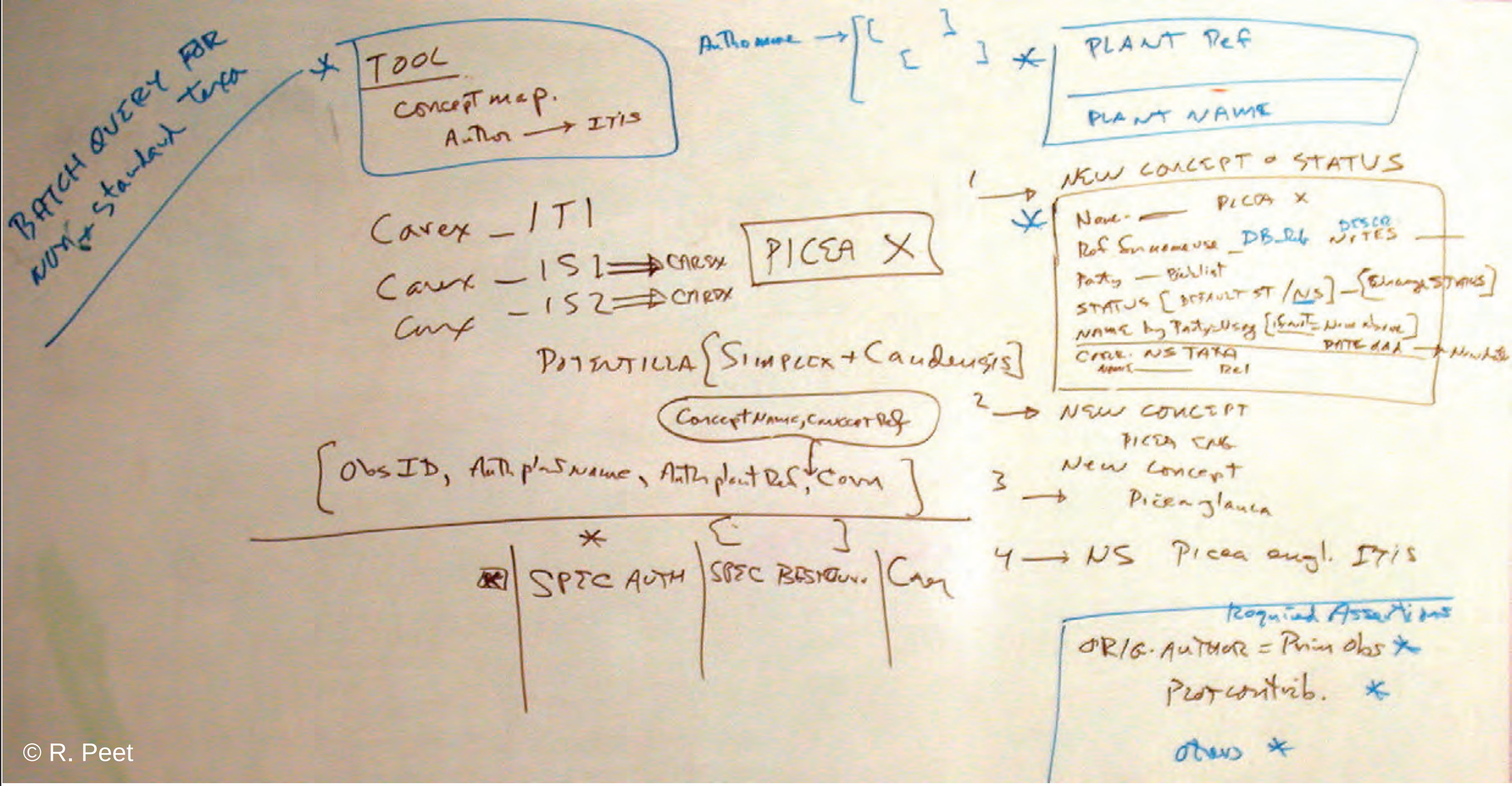

Fig. 3. White board at the US National Center for Ecological Analysis and Synthesis, February 2002, documenting the efforts of the international VegBank Design Committee to model inclusion of taxon concepts.

problem (e.g. Jansen & Dengler 2010, Peet et al. 2012) and led to potential database solutions (e.g. Franz et al. 2008, Franz & Peet 2009; Fig. 3), but such solutions require that large databases maintain information on the relationships among various taxon concepts. This information has, for the most part, not been consistently reported by the taxonomic community or embraced by that community as a disciplinary imperative. Nonetheless, vegetation scientists, and ecologists and biologists more generally, need to be able to unambiguously match and integrate organism names from a myriad of sources. We should partner with other organizations to bring about creation of the needed infrastructure.

4. Repeatability

The wealth of new analytical and statistical techniques used in the ecological community has led to wide use of R packages that assure that methods are easily and consistently used. The need for such tools is particularly strong in the case of extraction and analysis of data from large and complex databases. Recognizing this, both BIEN and sPlot have created R packages to facilitate data query and extraction. The next step is likely to be broader availability of work flows that automate certain kinds of analysis and that span multiple data sources (see Ludäscher et al. 2006, Reichman et al. 2011). IAVS could serve the community well by providing lists of critical and validated code (particularly R packages), and by advocating and documenting automated workflows for use by our community and others.

5. Insufficient data

Compilation of large databases also serves to reveal where data are sparse and more are needed. Examination of the distribution of the >1.1 million plots in sPlot reveals vast areas where it appears very few plots are available, such as Northern Africa and Siberia. In addition, certain types of data are limited in availability. For example, although several authors have advocated collection of data containing observations across multiple scales (e.g. Stohlgren et al. 1995, Peet et al. 1998, 2014, Dengler 2009), very few such datasets are available (but see Fig. 4). In addition, although the importance of long-term studies is broadly appreciated in vegetation science, few public archives contain plot data spanning long periods. The most notable exceptions include the system of plots maintained by the Center for Tropical Forest Science of the Smithsonian institution, with 63 large plots distributed over 24 countries and containing over 6 million trees representing over 10,000 species (http://www.forestgeo.si.edu/) and the US Forest Service Forest Inventory and Analysis plots (Gray et al. 2012). Another emerging example is the US National Ecological Observatory Network (NEON), which promises to provide at least 30 years of data for all taxa of plants (and many other organisms) across a geographically broad network (Kao et al. 2012). IAVS could serve the community by compiling information on what we view to be the most critical data needs, and advocating collection of such data to our membership and to various funding agencies.

6. Open culture

The number of journals and funding agencies that mandate open access to data has increased greatly over the last decade. Nonetheless, professional societies need to push harder for data to be open. IAVS could participate in the establishment and curation of archives of open data. Of course, this will require consistent policies and protocols with respect

Fig. 4. Collaborative data collection on a multi-scale permanent plot on Öland by Eddy van der Maarel, Robert Peet and Marijke van der Maarel, June 1985.

to such issues as managing intellectual property and the need for confidentiality to protect rare species and critical sites. Moreover, the need for greater openness goes beyond data. For example, sharing of computer code is not limited to R packages and should be mandated for most publications to ensure repeatability (Barnes 2010, Rocchini & Neteler 2012). At a minimum, IAVS should adopt and promote a set of best practices for openness with respect to data and code.

Time to think bigger!

Vegetation science lags behind other disciplines in the adoption of data protocols and the creation and maintenance of cyberinfrastructure. Genetic sequence data from around the world are nearly all in GenBank or equally accessible databases. Scientists do not hesitate to place their sequence data in such archives, or to extract data to support their on-going research. Other professional societies, such as the American Chemical Society, have standard repositories for critical types of data. Another model is TAIR, a comprehensive compilation of data from research on *Arabidopsis* that is curated and maintained by staff and to which research institutions subscribe so that their scientists have access (Huala et al. 2001). Perhaps the closest thing to a global vegetation science database that is both well populated and frequently used is TRY, the database for plant traits (Kattge et al. 2011), though some might question rules for access to TRY data.

There are at least three important ways in which vegetation data systems lag behind other disciplines. First, our data systems are rarely global, but more typically serve a subset of the world and often in different ways such that data from different systems do not integrate well. Second, participation in and curation of data systems is not yet a widely-embraced value in our field. We need for vegetation scientists to routinely deposit vegetation data in large repositories, and we need mechanisms to assure that the data are well curated so as to maintain high quality. Finally, many of our data systems are fragile in that they were created with short-term funding, and in most cases there is no clear vision for how to maintain these systems into the future and upgrade them as software requirements evolve.

Potential long-term roles for IAVS

Professional societies once served almost exclusively to publish journals and host meetings. Although both of these roles remain important, there are many other ways that professional societies could and should advance their fields. I have already observed many roles IAVS could play in the future to sustain and advance the growing role of data in our field. Among these are providing access to critical tools, hosting and curation of global data resources so as to assure their quality and longevity, providing mandates for archiving data and workflows, improving guidance and mechanisms for managing intellectual property and confidentiality, and providing advocacy for international programs for both one-time and long-term data collection following IAVS mandated protocols and with an emphasis on recognized data needs. There are many other possibilities.

How should we start? I have several suggestions. First, we need to collectively identify the needs and opportunities. Essentially, this is a call for a Vision Statement. I recommend that a special committee be created for this process. Second, we need to find ways to train a new generation of vegetation scientists with respect to recommended practices. We could sponsor workshops and we could develop materials for both undergraduate and graduate curricula. Finally, we need to develop a sustainable plan to accomplish these goals. In short, we need to develop a Business Plan. We should at least consider funding models from other fields. Certainly there is some potential in the Phoenix Bioinformatics Model followed by TAIR and BioCyc (e.g. Huala et al. 2001) where the user community pays a fee, which would generally be lower than if they were to do the work themselves. Finally, we should be aware that other professional societies face similar challenges and we should look for opportunities to collaborate with these other organizations so as to share the costs and enhance the benefits.

Acknowledgements

I cannot begin to recognize all the many friends, colleagues, students and family members who have supported my career and my participation in IAVS. However, two individuals stand out as having strongly shaped my goals, my career and my participation in IAVS. These are Robert Whittaker and Eddy van der Maarel. To them, and the many others, I am eternally grateful.

References

Anderson, M., P. Bourgeron, M.T. Beyer, R. Crawford, L. Engelking, D. Faber-Langendoen, M. Gallyoun, K. Goodin, D.H. Grossman, (…) & A.S. Weakley, A.S. 1998. International Classification of Ecological Communities: Terrestrial Vegetation of the United States. Volume II. The National Vegetation Classification System: List of Types. Arlington, VA: The Nature Conservancy.

Barnes, N. 2010. Publish your computer code: it is good enough. Nature 467: 753.

Barbour M.G., D.C. Glenn-Lewin, O.L. Loucks. 2000. Progress towards North American vegetation classification at physiognomic and floristic levels. In P.S. White, L. Mucina, J. Lepš (eds). Vegetation science in retrospect and perspective, Proceedings of the 41st IAVS symposium. Opulus, Grangärde, pp. 111–114.

Berendsohn, W.G. 1995. The concept of "potential taxa" in databases. Taxon 44: 207–212

Boyle, B., N. Hopkins, Z. Lu, J.A.R. Garay, D. Mozzherin, T. Rees, N. Matasci, M.L. Narro, W.H. Piel, (…) & B.J. Enquist. 2013. The Taxonomic Name Resolution Service: an online tool for automated standardization of plant names. BMC Bioinformatics 14: 16.

Chytrý M., S.M. Hennekens, B. Jiménez-Alfaro, I. Knollová, J. Dengler, F. Jansen, F. Landucci, J.H.J. Schaminée, S. Aćić, (…) & S. Yamalov. 2016. European Vegetation Archive (EVA): an integrated database of European vegetation plots Applied Vegetation Science 19: 173–80.

De Cáceres, M., M. Chytrý, E. Agrillo, F. Attorre, Z. Botta-Dukát, J. Capelo, B. Czúcz, J. Dengler, J. Ewald, (…) & S.K. Wiser. 2015. A comparative framework for broad-scale plot-based vegetation classification. Applied Vegetation Science 18: 543–560.

Dengler, J. 2009. A flexible multi-scale approach for standardised recording of plant species richness patterns. Ecological Indicators 9: 1169–1178.

Dengler, J., H. Bruelheide, O. Purschke, M. Chytrý, F. Jansen, S.M. Hennekens, U. Jandt, B.Jiménez-Alfaro, J. Kattge & V. Pillar, V. 2014. sPlot – the new global vegetation-plot database for addressing trait–environment relationships across the world's biomes. In: Mucina, L., Price, J.N. & Kalwij, J.M. (eds.). Biodiversity and vegetation: patterns, processes, conservation, Kwongan Foundation, Perth, Australia.

Dengler, J., F. Jansen, F. Glöckler, R.K. Peet, M. De Cáceres, M. Chytrý, J. Ewald, J. Oldeland, M. Finckh, (…) & N. Spencer. 2011. The Global Index of Vegetation-Plot Databases (GIVD): a new resource for vegetation science. Journal of Vegetation Science 22: 582–597.

Duke, C.S. 2006. Data: share and share alike. Frontiers in Ecology and the Environment 4: 395–448.

Enquist, B.J., R. Condit, R.K. Peet, M. Schildhauer & B.M. Thiers. 2016. Cyberinfrastructure for an integrated botanical information network to investigate the ecological impacts of global climate change on plant biodiversity. PeerJ Preprints 4: e2615v2 https://doi.org/10.7287/peerj.preprints.2615v2

Faber-Langendoen, D., K. Baldwin, R.K. Peet, D. Meidinger, E. Muldavin, T. Keeler-Wolf & C. Josse. 2017. The EcoVeg Approach in the Americas: U.S., Canadian and International Vegetation Classifications. Phytocoenologia, in press.

FGDC (Federal Geographic Data Committee). 2008. National Vegetation Classification standard, version 2. FGDC-STD-005-2008. Vegetation Subcommittee, U.S. Geological Survey, Reston, Virginia, USA.

Franz, N.M., R.K. Peet & A.S. Weakley. 2008. On the use of taxonomic concepts in support of biodiversity research and taxonomy. In: Wheeler, Q. D., Ed., The New Taxonomy. pp. 63–86. CRC Press, Boca Raton, FL.

Franz, N.M. & R.K. Peet. 2009. Perspectives: towards a language for mapping relationships among taxonomic concepts. Systematics and Biodiversity 7: 5–20.

Gauch, H.G. Jr. 1973. The Cornell ecology program series. Bulletin of the Ecological Society of America 54(3): 10–11.

Gauch, H.G. Jr. 1982. Multivariate analysis in community ecology. Cambridge University Press, Cambridge.

Gray, A.N., T.J. Brandeis, J.D. Shaw, W.H. McWilliams & P.D. Miles. Forest Inventory and Analysis Database of the United States of America (FIA). Biodiversity and Ecology 4: 225–231.

Guralnick, R.P., Wieczorek, J., Beaman, R. & Hijmans, R.J. 2006. BioGeomancer: automated georeferencing to map the world's biodiversity data. PLoS Biology 4: e381.

Hennekens, S.M. 1995. TURBO(VEG). Software package for input, processing, and presentation of phytosociological data. User's guide. Instituut voor Bos en Natuur, Wageningen and Unit of Vegetation Science, University of Lancaster, Lancaster.

Hennekens, S.M. & J.H.J. Schaminée. 2001. TURBOVEG, a comprehensive database management system for vegetation data. Journal of Vegetation Science 12: 589–591.

Hill, M.O. 1979. TWINSPAN: A Fortran program for arranging multivariate data in an ordered two-way table by classification of the individuals and attributes. Cornell University, New York, USA

Huala, E., A.W. Dickerman, M. Garcia-Hernandez, D. Weems, L. Reiser, F. LaFond, D. Hanley, D. Kiphart, M. Zhuang, (…) & S.Y. Rhee. 2001. The Arabidopsis Information Resource (TAIR): a comprehensive database and web-based information retrieval, analysis, and visualization system for a model plant. Nucleic Acids Research 29: 102–105.

Jansen, F. & Dengler, J. 2010. Plant names in vegetation databases – a neglected source of bias. Journal of Vegetation Science 21: 1179–1186.

Jennings, M.D., D. Faber-Langendoen, O.L. Loucks, R.K. Peet & D. Roberts. 2009. Standards for associations and alliances of the U.S. National Vegetation Classification. Ecological Monographs 79: 173–199.

Jones, M.B., M. Schildhauer, O. J. Reichman & S. Bowers. 2006. The new bioinformatics: Integrating ecological data from the gene to the biosphere. Annual Reviews of Ecology, Evolution and Systematics 37: 519–544.

Kao, R. H., C. M. Gibson, R. E. Gallery, C. L. Meier, D. T. Barnett, K. M. Docherty, K. K. Blevins, P. D. Travers, A. Azuaje, (…) & D. Schimel. 2012. NEON terrestrial field observations: designing continental-scale, standardized sampling. Ecosphere 3(12): 115, 6p.

Kattge, J., S. Díaz, S. Lavoral, I.C. Prentice, P. Leadley, G. Bönisch, E. Garnier, M. Westoby, P.B. Reich, (…) & C. Wirth. 2011. TRY – a global database of plant traits. Global Change Biology 17: 2905–2935.

Kelling, S., W.M. Hochachka, D. Fink, M. Riedewald, R. Caruana, G. Ballard & G. Hoper. 2009. Data-intensive Science: A New Paradigm for Biodiversity Studies. BioScience 59: 613–620.

Ludäscher, B., I. Altintas, C. Berkley, D. Higgins, E. Jaeger, M. Jones, E. Lee, J. Tao & Y. Zhao. 2006. Scientific workflow management and the Kepler system. Concurrency and Computation: Practice and Experience 18: 1039–1065.

Madin, J.S., S. Bowerts, M.P. Schidhauer & M.B. Jones. 2008. Advancing ecological research with ontologies. Trends in Ecology and Evolution 23: 159–168.

May, R.M. 1986. The search for patterns in the balance of nature: advances and retreats. Ecology 67: 1115–1126.

Michener, W.K. & M.B. Jones. 2012. Ecoinformatics: supporting ecology as a data-intensive science. Trends in Ecology and Evolution 27: 85–93.

Mucina, L., G. Grabherr, G. & T. Ellmauer, T. (eds.) 1993. Die Pflanzengesellschaften Österreichs, Gustav Fischer, Jena.

Mucina, L., J.S. Rodwell, J.H.J. Schaminée, & H. Dierschke. 1993. European Vegetation Survey: current state of some national programs. Journal of Vegetation Science 4: 429–438.

Mucina, L., J. H J. Schaminée, & J. S. Rodwell. 2000. Common data standards for recording relevés in field survey for vegetation classification. Journal of Vegetation Science 11: 769–772.

Mucina, L., H. Bültmann, K. Dierßen, J.-P. Theurillat, T. Raus, A. Čarni, K. Šumberová, W. Willner, J. Dengler, (…) & S.M. Hennekens. 2016. Vegetation of Europe: Hierarchical floristic classification system of vascular plant, bryophyte, lichen, and algal communities. Journal of Applied Vegetation Science 19(Suppl. 1): 3–264.

Orlóci, L. 1978. Multivariate analysis in vegetation research. 2nd Edition. Junk, The Hague

Peet, R.K. 1998. ESA journals: evolution and revolution. Bulletin of the Ecological Society of America 79(3): 177–181.

Peet, R.K., M.T. Lee, M.D. Jennings & D. Faber-Langendoen, D. 2012. VegBank – a permanent, open-access archive for vegetation-plot data. Biodiversity and Ecology 4: 233–241.

Peet, R.K., K.A. Palmquist & S.M. Tessel. 2014. Herbaceous layer species richness of southeastern forests and woodlands: Patterns and causes. In: Gilliam, F.S. (ed.), The Herbaceous Layer in Forests of Eastern North America, Second edition. pp. 255–276. Oxford University Press, Oxford, UK.

Peet, R.K., T.R. Wentworth & P.S. White. 1998. A flexible, multipurpose method for recording vegetation composition and structure. Castanea 63: 262–274.

Pielou, E.C. 1984. The Interpretation of Ecological Data. Wiley, New York.

Purschke, O., J. Dengler, H. Bruelheide, M. Chytrý, F. Jansen, S. Hennekens, U. Jandt, B. Jiménez-Alfaro, J. Kattge, (…) & M. Winter. 2015. sPlot – the new global vegetation-plot database for addressing trait-environment relationships across the world's biomes. EGU General Assembly 2015. http://adsabs.harvard.edu/abs/2015EGUGA..1715724P.

Reichman, O.J., M.B. Jones & M.P. Schildhauer. 2011. Challenges and opportunities of open data in ecology. Science 331, 703–705

Rocchini, D. & M. Neteler. 2012. Let the four freedoms paradigm apply to ecology. Trends in Ecology and Evolution 27: 310–311.

Rodwell, J.S., S. Pignatti, L. Mucina & J.H.J. Schaminée. 1995. European Vegetation Survey: update on progress. Journal of Vegetation Science 6: 759–762.

Rodwell, J.S. 1991–2000. British Plant Communities, Vols. 1–5. Cambridge University Press, Cambridge.

Schaminée, J.H.J., P.W.F.M. Hommel, A.H.F. Stortelder, E.J. Weeda & V. Westhoff. 1995–1999. De Vegetatie van Nederland, Vols. 1–5. Opulus Press, Uppsala/Leiden.

Stohlgren, T. J., M.B. Falkner & L.D. Schell. 1995. A modified-Whittaker nested vegetation sampling method. Vegetatio 117: 113–121.

van der Maarel, E. 1971. Bericht über die arbeitsgruppe für datenverarbeitung der internationalen vereinigung für vegetationskunde, 1969–1970. Vegetatio 22: 269–273

van der Maarel, E. 1974. The Working Group for data-processing of the International Society for Plant Geography and Ecology in 1972–1973. Vegetatio 29: 63–67.

van der Maarel, E. 1980. Introduction. Vegetatio 42: VII–VIII.

van der Maarel, E. 1990. The Journal of Vegetation Science: a journal for all vegetation scientists. Journal of Vegetation Science 1: 1–4.

Van Kleunen, M., W. Dawson, F. Essl, J. Pergl, M. Winter, E. Weber, H. Kreft, P. Weigelt, J. Kartesz, M. Nishino, (…) & P. Pyšek. 2015. Global exchange and accumulation of non-native plants. Nature 525: 100–103.

Walker, D. A., A.L. Breen, L.A. Druckenmiller, L.W. Wirth, W. Fisher, M.K. Raynolds, J. Šibík, M.D. Walker, M. D., (…) & D. Zona. 2016. The Alaska Arctic Vegetation Archive (AVA-AK). Phytocoenologia 46: 221–229.

Walker D.A., F.J.A. Daniëls, N.V. Matveyeva, J. Šibík, M.D. Walker, A.L. Breen, L.A. Druckenmiller, M.K. Raynolds, H. Bültmann, (…) & L.M. Wirth. 2017. Circumpolar Arctic Vegetation Classification. Phytocoenologia, in press.

Wiser, S.K. 2016. Achievements and challenges in the integration, reuse and synthesis of vegetation plot data. Journal of Vegetation Science 27: 868–879.

Wiser, S.K., N. Spencer, M. De Cáceres, M. Kleikamp, B. Boyle & R.K. Peet. 2011. Veg-X – An international exchange standard for plot-based vegetation data. Journal of Vegetation Science 22: 598–609.