

PLANT NAME RESOURCES: BUILDING BRIDGES WITH USERS

Alan Paton,^{✉1} Robert Allkin,¹ Irina Belyaeva,¹ Elizabeth Dauncey,¹ Rafaël Govaerts,¹ Sarah Edwards,¹ Jason Irving,¹ Christine Leon,¹ and Eimear Nic Lughadha¹

Plant names are the key to communicating and managing information about plants. This paper considers how providers of high quality technical plant name information can better meet the requirements non-botanical audiences who also rely on plant names for elements of their work. The International Plant Name Index, World Checklist of Selected Plant Families and The Plant List are used as examples to illustrate the strengths and weaknesses of plant name resources from a non-expert user's perspective. The above resources can be thought of as botanists pushing data at audiences. Without closer engagement with users, however, there is a limit to their relevance and impact. The need to cover common names is a frequent criticism of existing resources. The Medicinal Plant Names Services (MPNS, www.kew.org/mpns) is an example of how plant name resources can be adapted to better address the needs of a non-botanical audience. Some of the major challenges are outlined and solutions suggested.

INTRODUCTION

Plant names are the means by which we find information about plants (Paton *et al.* 2008; Patterson *et al.* 2010; Allkin, 2014a). However, plants have on average three different scientific names¹ each: roughly 370, 000 vascular plant species have almost one million names at species level. Commonly used plants have many more names; medicinal plants,

¹ Royal Botanic Gardens Kew, Richmond, Surrey, TW93AE, UK; Email: a.paton@kew.org.

[✉] Corresponding author.

This is the author manuscript of the chapter published in *Botanists of the twenty-first century: roles, challenges and opportunities*, edited by Noëline R. Rakotoarisoa, Stephen Blackmore, and Bernard Riera.

NLM Citation: Paton A, Allkin R, Belyaeva I, et al. PLANT NAME RESOURCES: BUILDING BRIDGES WITH USERS. In: Rakotoarisoa NR, Blackmore S, Riera B, editors. *Botanists of the twenty-first century: roles, challenges and opportunities* [Select Chapters]. Paris, France: United Nations Educational, Scientific and Cultural Organisation (UNESCO); 2016.

This chapter is open access under a Creative Commons Attribution-ShareAlike license (CC-BY-SA).

¹ Scientific plant names are routinely referred to as 'Latin names', however, pharmaceutical names used in many pharmacopoeias are also written in Latin and referred to by practitioners,

for example, have on average 12 scientific names² at species level and some have many more. Non-scientific names are often used by professionals in disciplines other than botany. Hence a single species may be referred to by multiple scientific and non-scientific names causing confusion and ambiguity - with potentially serious consequences. For example, inappropriate plants have been used in the production of herbal medicinal products, resulting in deleterious health effects to the consumer.

How can botanists meet the needs of diverse user communities for easy, intelligible access to, and use of, plant names? Part of the solution is technological, for example the provision of systems which can facilitate linking scientific plant names to other names and information resources. However, the greater barriers are sociological. What do non-botanical audiences want? How can botanists discover their needs? How can we build relationships with audiences focused on plant uses and help them to use plant names effectively, accurately and unambiguously? Should we prioritise their needs over those of the botanical community?

This paper presents the Royal Botanic Gardens, Kew's current plant name resources and attempts to outline their strengths and weaknesses from a non-expert user's perspective. We then explore how the needs of these users can be better met. We use the Medicinal Plant Names Services project as an example of how we are addressing some of the challenges of effective delivery of plant name information to a non-botanical audience.

Plant Names Resources

In 1882, Charles Darwin contacted Joseph Hooker, the then Director of Royal Botanic Gardens at Kew, with an offer to provide in his will "about £250 a year for 4 or 5 years for the formulation of a perfect m.s. catalogue of all known plants.". Darwin had used Steudel's Nomenclator (Steudel 1840) for his work and recognized the value of such a tool. The task fell to Daydon Jackson who realised that a catalogue of all known plants would quickly go out of date. He suggested that a list of plants described, with their bibliographic details, would be of immense value, and could be maintained (Jackson 1887, 1924). The first volume of *Index Kewensis* was produced in 1895.

The origins of *Index Kewensis* illustrate some important points of equal relevance today. Darwin had a particular requirement and interest. Jackson and Hooker discussed with Darwin how his needs might be best addressed given the resources at hand. Funding was identified to enable the project to begin and a host organisation was prepared to support its further development. In other words, the strengths and core activities of an institution were focused to develop a new product to meet Darwin's need and those of many others.

manufacturers and health regulators as 'Latin names'. To avoid confusion, we use the terms 'scientific name' and 'pharmaceutical name' in this paper.

² Data from Medicinal Plant Names Services resource, see later for details.

In 2000 the *International Plant Names Index* (IPNI) (www.ipni.org) was launched as an on-line collaboration between *Index Kewensis* (The Royal Botanic Gardens, Kew), the Gray Herbarium Card Index (Harvard University Herbaria) and the Australian Plant Name Index (Centre for Plant Biodiversity Research, Canberra), initially funded by the US National Science Foundation and the US Geological Survey (Croft et al. 1999). IPNI is a database of the names and associated basic bibliographic details of seed plants, ferns and lycophytes. Its goal is to eliminate the need for repeated reference to primary sources for basic bibliographic information about plant names. The data are freely available and are gradually being standardised and checked. IPNI is a dynamic resource, depending on contributions by many members of the botanical community. IPNI aims to be comprehensive for all plant names within its scope. It includes over 1.6 million plant names and the editorial team keep it up to date, adding new names and checking, linking and standardising earlier records. Although IPNI is aimed at a taxonomic audience, it does not provide taxonomic synonymy, but focuses purely on nomenclature: it is a list of names rather than of plant taxa. The difference between a name and a taxon concept - which includes all the names relevant to a particular taxon - is a common source of confusion for the more general user and even many botanists!

The task of creating a list of plants and assigning all possible synonyms to each is huge and, as Jackson rightly recognised, any such list is soon outdated as evidence accrues from further studies. Ultimately, of course, this is the work of the entire botanical community and requires the support of multiple “curators”. The urgent need for a working synonymised list of plants was articulated as Target 1 of the Global Strategy for Plant Conservation (GSPC) in 2002. Such prominence in the GSPC acknowledges the importance of being able to communicate unambiguously about plants, and that names used for a particular plant taxon are key to finding information relevant to its conservation (Paton *et al.* 2008). We highlight some of issues faced by non specialists using two examples of these systems: the *World Checklist of Selected Plant Families* and *The Plant List*.

The *World Checklist of Selected Plant Families* (WCSP, www.kew.org/wcsp) covers all species in 174 families of vascular plants. It is complete for monocots and for other large, important plant families such as the coffee family (Rubiaceae), the mints (Lamiaceae), spurges (Euphorbiaceae) and myrtles and eucalypts (Myrtaceae). WCSP is compiled from the botanical literature and then reviewed by experts. One of its strengths is that it is supported by over 160 collaborators around the world, who contribute to and review the data. It provides full synonymy, distribution information to country level enabling geographically based searches, information on life form, bibliographic references to taxonomic concepts, and links to further information. However, coverage is incomplete (126,400 species), and it may be unclear to non-specialist users whether their plant of interest is treated or not. WCSP is aimed at botanists and horticulturists using scientific names.

The Plant List (*TPL*, www.theplantlist.org) is a collaborative venture coordinated by the Royal Botanic Gardens, Kew and Missouri Botanical Garden, and relies on collaborators

managing significant taxonomic data resources. The approach adopted was to merge into a single consistent database the best of the taxonomic information available in digital form through a defined and automated process. Accepted names and synonymy relationships from global checklist datasets were augmented by additional names and synonymy relationships drawn from floristic (regional and national) datasets. An automated process was used to compare the taxonomic judgements expressed within the diverse datasets and conflicts or inconsistencies were resolved using suites of logical rules designed to mimic the thought processes a taxonomist might use to decide between two divergent accounts. Species names not incorporated as a result of this process were added from IPNI, rendering the TPL comprehensive for plant names. Final checks and adjustments enhanced logical consistency.

The Plant List met its objective of providing a 'best effort' list, responding to the demand for a working list and stimulating further efforts. The strengths of TPL (Allkin 2014b) are that it is comprehensive, indicates synonymy based upon the contributing data sources and uses a star rating to indicate the relative confidence in the status of each name record. The user interface is simple and attractive with links to data sources and other information. Nonetheless, TPL is far from perfect and represents work in progress. Around 25% of names in TPL are 'unresolved' indicating that the source data sets included no information on their taxonomic status or conflicted with one another. Several authoritative sources of taxonomic opinion for large groups or regions were not included in the current version of TPL simply due to constraints on time and resources. TPL is static: it is neither edited directly nor updated regularly from the original data sources. Feedback and corrections pertaining to records in TPL are passed to the source database for consideration. If accepted by the source database they may be incorporated in a future version of TPL. This results in TPL data becoming out of date. Unlike WCSP, TPL contains no geographical information.

Despite limitations in the quality of data it provides, TPL attracts far more use than IPNI or WCSP from a much wider set of users (Table 1). It seems the audiences for IPNI and WCSP are more sophisticated users, comfortable with the more complex interfaces and interpreting high quality technical information. Feedback from users suggests that the audiences of TPL are attracted by its comprehensive nature, its ease of use and the external linkages to other information resources. These linkages also contribute to TPL's prominent position in the results of web search engines. One of most frequent criticisms received by the editors of these plant name resources are that they do not deal with common names and lack validated images.

Deciding how best to interpret common names which are used to refer to more than one plant is a problem for many users. The name 'fang ji', for example, is a noteworthy case in point. Owing to certain shared clinical attributes, this name was widely used within Traditional Chinese Medicine (TCM) to refer to the root of either *Stephania tetandra* S.Moore or *Aristolochia fangchi* Y.C.Wu. This dual usage arose due to an understanding within TCM that these species could be used interchangeably to treat certain medical conditions. Their chemistry, however, is quite different, with *A. fangchi* (like all species in

this genus) containing nephrotoxic compounds called aristolochic acids, together with carcinogens. In the early 1990s, unaware of these toxicity differences and indeed probably also that this common name could refer to more than one species, a Belgian slimming company included ‘fang ji’ in one of its slimming products (unconnected with TCM), with fatal consequences. The ‘fang ji’ included was the *Aristolochia* species and not the *Stephania*; the latter is devoid of these toxic compounds. The outcome was 115 patients with kidney failure, of which 46 also developed urinary carcinomas. At least four patients died (Gokmen and Lord, 2012; Nortier *et al.*, 2000; Vanherweghem *et al.*, 1993). Had there been greater awareness of the potential confusion over the identity of ‘fang ji’, these tragic outcomes might have been avoided.

Table 1. Usage statistics the plant name resources calculated from Google analytics. Number of visits (sessions) and number of unique visitors calculated from 1 January 2015 to 31 December 2015.

Resource	Number of visits	Number of unique users
The Plant List	2,726,096	1,337,771
IPNI	589,733	213,683
World Checklist	269,894	112,838

METHODS: DETERMINING THE NEEDS OF USERS

To varying degrees IPNI, WCSP and TPL are aimed at plant scientists, some particularly at systematists. However, high quality plant names underpin everyone’s ability to communicate and discover information about plants. How can the information in these resources be made more accessible and useful to a broader audience? To answer this question it is vital to understand users’ requirements: what information do users need, and how is it to be used?

Kew’s Medicinal Plant Names Services (MPNS) (www.kew.org/mpns) is a project designed to build a global resource for medicinal plant names, enabling health professionals and researchers to access information about plants relevant to pharmacological research, health regulation and traditional medicine. Although the resource has relevance for a much wider audience, for example conservationists, MPNS deliberately focused on a particular group of related users so as to determine more precisely the requirements of this user community. From the outset, identification of and interaction with the user landscape was prioritised, with a significant budget for this activity, as well as time and resources to respond to lessons learnt both in terms of the design and functionality of the services, and the scope of the underlying resource.

A range of different approaches including conferences, questionnaires, invited workshops, one to one meetings and planning meetings were selected as means of engaging with prospective users to discover their requirements and priorities. Our involvement with these communities of users gradually intensified as we have understood more clearly how to engage and as those communities have come to recognise our work. Professional meetings provide opportunities to communicate and disseminate the services that we

offer, explain their relevance and hear suggestions. Workshops target particular audience segments with which we can work to define the benefits to them of more efficient and more rigorous use of plant names. Workshops enable us to test ideas, observe users interacting with proposed systems and obtain feedback to drive improvements. One to One meetings allow us to learn of local functional requirements and priorities.

RESULTS: MEDICINAL PLANT NAMES SERVICES

Extended and diversified engagement with users has made us aware of major requirements such as the absolute necessity of covering pharmaceutical, trade and common names (and how these map onto the botanical nomenclature) as well as the ability of users to find scientific names exactly as they are used within the medicinal literature and regulations (however misspelt or misused they may be). We also better understand how plant names are captured in users' own systems. In many cases their existing data structures are unable to deal appropriately with plant names. For example, they may not be able to deal with multiple names for a single species (synonymy) or appreciate that the same binomial may refer to more than one species (homonymy). Many users request training as to how best to deal with the complexities of plant names.

To meet the identified demands, MPNS has designed and built several outputs in the form of information services. An open access portal provides for simple manual search of scientific, pharmaceutical (pharmacopoeial) and common names (<http://apps.kew.org/mpns-portal/>). A search using 'Fang Ji' returns the currently accepted scientific names of both species to which 'Fang Ji' may be applied, along with the alternative scientific names for these species, other relevant non-scientific names and references in which those names appear. Further information tabs take the user to links to external digital resources and allow the user to search these using either the accepted scientific name alone, or all scientific synonyms of that plant simultaneously. As might be anticipated the latter searches return far more data records since many will be stored under older synonyms.

MPNS will offer other services designed to meet more complex needs. These include:

1. **Name Validation:** checking lists of plant names held by other organisations and individuals. MPNS can correct spellings, propose updates to taxonomy and nomenclature, and enrich such lists with all known synonyms for each plant. This enables detection of single plants appearing repeatedly under alternative names, and can embed unique digital identifiers into client databases to facilitate maintenance of data and future updates.
2. **Web Services:** connecting user IT systems to MPNS electronically enabling validation of names as they are entered, the refreshing of client databases as plant names change, building comprehensive names indexes, and intelligent data mining.
3. **Vocabulary Control:** providing authoritative reference lists, ontologies, terminological controls and data subsets for use in other organisation's information systems.

4. Harmonisation: mapping plant lists onto plants cited in legislation or by other organisations or publications so as to show overlaps, detect gaps and enhance communication.
5. Consultancies: providing expert advice in the use and interpretation of medicinal plant names, as well as in devising workflows to capture and store scientific plant names appropriately and designing database structures to manage these names.
6. Training: specialist training courses for people working with medicinal plant names to enable safe and efficient working practices. Courses at Kew or on-site for larger clients.

Access to the portal is free, but there are charges for some of the more complex, value-added services, generating revenue to help cover the costs of maintaining and updating the core taxonomic and nomenclatural data as well as the costs of providing the services themselves.

DISCUSSION: PLANT NAMES RESOURCES IN THE 21ST CENTURY

MPNS exemplifies a way ahead for botanical resources. It demonstrates the relevance and importance of taxonomic research but provides mechanisms for non-specialist users to access that information. Three major challenges are inherent in the approach taken by MPNS: technical, sociological and resourcing.

Technical Challenges

Technical innovations are required to maximise the potential reuse of data. Key approaches include attaching persistent identifiers to elements of knowledge such as names and taxonomic concepts (the accepted name and its set of synonyms). Persistent identifiers allow users always to refer to a particular record unambiguously and facilitate transfer of information from one computer-based system to another. Persistent identifiers for scientific names have been implemented by IPNI and others, but are lacking for taxonomic concepts which would be more useful to the MPNS audience. The MPNS audience are interested in links to other data resources. Registries of information content would make potentially available resources more easy to discover and facilitate linkage.

Sociological Challenges

Unlike traditional botanical resources, the emphasis of MPNS is not on providing the correct name which should be used for a particular species, although that information is indicated. Rather the focus is on allowing users to appreciate what names might have been applied to the plant they are interested in, and using that data to mine other resources or manage their own data more effectively. This is a change of mind-set for many botanists and involves developing new practices, such as maintaining a record of misspellings and inappropriate use of names in order to interpret data from resources with less reliable plant name data. The nomenclatural purist would seek to “correct” those

errors and hide them. In the real world these names are in use. The user group offers opportunities since they have their own information on plant names, and hold data which may be missing from MPNS. Viewing users and providers as one community who can work together will lead to stronger relationships and a higher quality product.

Although MPNS does not seek to impose the use of particular names, through use of MPNS over time, different user groups may improve their understanding of plant names and perhaps move towards an increasing consensus as to which names to use in preference to others. That such consensus may ultimately be reached voluntarily is indeed a goal worth working towards, since this may gradually introduce increased harmonisation of names reducing current confusion, for the benefit of all.

Working with the developers of professional data standards and legislation relevant to a particular user group is another way of disseminating the value of the MPNS resource to those working with medicinal plant names. MPNS has provided expert advice on the implementation of an ISO standard³ for the unique identification of substances used in medicinal products. By co-authoring the Herbal annex and providing controlled vocabularies for plant parts and plant names, MPNS is contributing to the ISO's requirement for consistency in the naming of substances, and will benefit from increased impact and sustainability for the MPNS resource and associated services.

Resource Challenges

There is a growing demand for open access to information. Currently many resources make data available under restrictive licences, hampering the flow of information. The EU Framework 7 project i-proBiosphere developed the Bouchout Declaration (<http://www.bouchoutdeclaration.org/declaration/>) as a set of principles for institutions and individuals to sign and support. These outline the need for open access and some of the technical requirements, mentioned above, necessary to create an 'open biodiversity knowledge management' system. The aim is to improve the availability of biodiversity information and its management. Adoption of these principles will greatly enhance the impact and utility of the information that botanists produce. Pressures for open access to data derive increasingly from Government and from funding bodies. The Wellcome Trust that supports MPNS has an explicit open access policy "*The Wellcome Trust therefore supports unrestricted access to the published output of research as a fundamental part of its charitable mission and a public benefit to be encouraged wherever possible*" (<http://www.wellcome.ac.uk/About-us/Policy/Policy-and-position-statements/WTD002766.htm>). Nevertheless The Wellcome Trust Programme Officers express equal concern for the sustainability of resources and services developed by MPNS.

³ ISO/DTS19844 Health Informatics – Identification of medicinal products – Implementation Guide for EN ISO 11238 Data elements and structures for the unique identification and exchange of regulated information on substances (2014)

That MPNS has been developing an explicit business model to sustain both data resources and the services built upon them may, at first glance, appear contrary to the principle of open access. Institutional support for such resources is critical, and in particular the technical structures, core data and business needs must be embedded within the strategic aims of the institution. As pressure increases on budgets for publically funded research in general, enhanced cost recovery even for partly publically funded institutions is vital. MPNS approaches this problem in several ways. The portal is open access whilst value added services are provided under licence which may therefore be charged for and the income used to help offset costs. The development of MPNS has been designed to be modular, enabling the scope of services or name data to increase as further funding from clients or grants becomes available. Cost savings can result from building a core institutional information architecture which can service different communities with appropriate, tailored data. Developing income-generating services for each new group requires close communication with the intended users to ensure a clear understanding of their specific needs and how the benefits of the service to the user outweigh the costs.

CONCLUSIONS

Plant name resources should serve a wider community than just the botanists and systematists that have previously been their creators and their audience. Additional users of plant names will differ in their requirements. Understanding these requirements is best achieved through collaboration between data providers and potential users, and it takes time to nurture these complex relationships. Plant names resources need to become more collaborative with providers, who understand the strengths and limitations of their data, reaching out to potential users, who may also have a role in improving or adding data. It is important to obtain clearly articulated user needs including how best to disseminate data for maximum benefit. The providers of plant name resources will need to be more business-like in their approach, understanding their costs, potential income streams and balancing the needs of open access and income generation through value added services or grants. Building bridges with users will ultimately increase the impact of plant name data across society.

ACKNOWLEDGEMENTS

MPNS is funded by the Wellcome Trust Grant #046369: “Plant Name Information Services. Work on the Bouchout Declaration was funded as part of pro-iBiosphere project (<http://www.pro-ibiosphere.eu/>) funded by the EU Framework Programme (FP7/2007-2013) under grant agreement N°312848.

REFERENCES

ALLKIN, R. (2014a) Communicating safely and effectively using plant names. Chapter 11 in *Traditional Medicine and Globalization*, Society of Ethnopharmacy, Maven (Kolkata) pp 111-125.

- ALLKIN R. What is in a name? The need for accurate scientific nomenclature for plants. *Journal of Ethnopharmacology*. 2014b;152:393–402. PubMed PMID: 24374235.
- CROFT J., CROSS N., HINCHCLIFFE S., LUGHADHA E. N., STEVENS P. F., WEST J. G., WHITBREAD G. Plant names for the 21st century: the International Plant Names Index, a distributed data source of general accessibility. *Taxon*. 1999.:317–324.
- GOKMEN M.LORD G. Aristolochic Acid Nephropathy: A disease that could be prevented through more careful regulation of herbal products. *BMJ*. 2012;344(E4000): 1–2. AND.
- INTERNATIONAL PLANT NAME INDEX (2014). Published on the Internet; www.ipni.org [accessed 1 October 2014].
- JACKSON B. D. The history of the compilation of the Index Kewensis. *Journal of the Royal Horticultural Society of London*. 1924;49:224–9.
- JACKSON B. D. 1887 the new 'Index of Plant Names'. *Journal of Botany, British and Foreign* 25: 66-71, 150-151.
- Medicinal Plant Names Services Portal (2014), Royal Botanic Gardens, Kew, accessed on 01/10/2014.
- NORTIER J.L., et al. Urothelial carcinoma associated with the use of a Chinese herb (*Aristolochia fangchi*). *N Engl J Med*. 2000;342:1686–1692. PubMed PMID: 10841870.
- PATON A. J., BRUMMITT N., GOVAERTS R., HARMAN K., HINCHCLIFFE S., ALLKIN B., LUGHADHA E. N. Towards Target 1 of the Global Strategy for Plant Conservation: a working list of all known plant species—progress and prospects. *Taxon*. 2008;57(2):602–611.
- PATON A. Biodiversity informatics and the plant conservation baseline. *Trends in Plant Science*. 2009;14(11):629–637. PubMed PMID: 19783196.
- PATTERSON D. J., COOPER J., KIRK P. M., PYLE R. L., REMSEN D. P. Names are key to the big new biology. *Trends in ecology & evolution*. 2010;25(12):686–691. PubMed PMID: 20961649.
- STEUDEL, E. G. (1840). *Nomenclator botanicus, seu: Synonymia plantarum universalis, enumerans ordine alphabetico nomina atque synonyma, tum generica tum specifica, et a Linnaeo et a recentioribus de re botanica scriptoribus plantis phanerogamis imposita* (Vol. 1). Typis et sumptibus JG Cottae.
- THE PLANT LIST (2014). Version 1.1. Published on the Internet; www.theplantlist.org/ (accessed 1 October 2014).
- VANHERWEGHEM J. L., TIELEMANS C., ABRAMOWICZ D., DEPIERREUX M., VANHAELEN-FASTRE R., VANHAELEN M., JADOUL M. Rapidly progressive interstitial renal fibrosis in young women: association with slimming regimen including Chinese herbs. *The Lancet*. 1993;341(8842):387–391. PubMed PMID: 8094166.
- WCSP (2014). *'World Checklist of Selected Plant Families*. Facilitated by the Royal Botanic Gardens, Kew. Published on the Internet; Available at: <http://apps.kew.org/wcsp/> Retrieved 1 October 2014.