Global Plant Name data online: their relevance to horticulture (Part 1)

Alan Paton, Rafaël Govaerts, Bob Allkin
Overview

• Brief overview of some global resources

• Focus on the issues these resources face

• Lessons learned
Examples

• International Plant Names Index
• World Checklist of Selected Plant Families
• Tropicos
• Catalogue of Life
• The Plant List

• Focus on species: none of these cover cultivar names
The International Plant Names Index (IPNI) is a database of the names and associated basic bibliographical details of seed plants, ferns and lycophytes.
1) International Plant Name Index

Catalogue ALL published scientific names for Vascular plants

Collaboration Kew, Harvard, Australia:
Index Kewensis started in 1885

Example use:
- Is this name valid?
- How should it be spelt?
- Where and when was it published?
- Are there homonyms?

Strengths:
- Complete
- 1.6 million names
- Editorial team: up-to-date
- And heavily curated

Weaknesses:
- Intended for taxonomists!
- No taxonomy
- Incomplete for infraspecifics
- Scope: not algae or mosses

www.ipni.org
Quick Search

This Checklist gives information on the accepted scientific names and synonyms of selected plant families. It allows you to search for all the scientific names of a particular plant, or the areas of the world in which it grows (distribution).

The checklist includes 151 Seed Plant families (View list of included families or reviewers). Different families are in different stages of review as indicated in the family list.

Search for a family, genus, or genus plus species e.g. *Amalia acuminata*

(enter family names in full and use the wildcard character (*) for partial matches on genus and species.)

To perform a more detailed search go to the Advanced Search Page.

To build your own checklist go to the Build a checklist Page.

The currency of science is citation so if you find this website useful and use it for your publications, please cite us. For more information about the checklist please contact Rafaël Govaerts.
2) World Checklist

Built at Kew: 150 systematists worldwide
Lists all taxa in the 173 families covered
Complete for monocotyledons
Links each species to a) all known synonyms  b) geography

Example use:
What is current name and ALL the synonyms?
Which publication accept/synonymise this name?
Where does species X grow?

Strengths:
- Links all synonyms
- Geography; lifeform
- Global and all data referenced
- High quality / peer review
- Links to other data

Weaknesses:
- Incomplete: 50% of plants
- Scientific names only
- No images or descriptions

www.kew.org/wcsp
Tropicos® was originally created for internal research but has since been made available to the world's scientific community. All of the nomenclatural, bibliographic, and specimen data accumulated in MBG's electronic databases during the past 25 years are publicly available here. This system has over 1.2 million scientific names and 4.0 million specimen records.

This map shows the country distribution of Tropicos specimen records that have country specified. Click the image to see a larger version and to view counts per country.
Tropicos

Strengths

• Curated
• MO research strengths
• Images; specimens
• Links out to other data sources

Weaknesses

• incomplete for plants
• Geographic bias in coverage
Catalogue of Life
www.catalogueoflife.org/
Catalogue of Life

Strengths

• Aims to cover all biodiversity
• A network of Global Species Databases
• Data curation by source data providers
• Full synonymy
• Coverage of all vascular plants

Weaknesses

• Some sources provide distribution and common names others do not
• Variable in data quality
The Plant List [www.theplantlist.org](http://www.theplantlist.org)

The Plant List is a working list of all known plant species. Version 1 aims to be comprehensive for species of Vascular plant (flowering plants, conifers, ferns and their allies) and of Bryophytes (mosses and liverworts).

Collaboration between the Royal Botanic Gardens, Kew and Missouri Botanical Garden enabled the creation of The Plant List by combining multiple checklist data sets held by these institutions and other collaborators.

The Plant List provides the Accepted Latin name for most species, with links to all Synonyms by which that species has been known. It also includes Unresolved names for which the contributing data sources did not contain sufficient evidence to decide whether they were Accepted or Synonyms.

Summary Statistics
The Plant List includes 1,040,426 scientific plant names of species rank. Of these 298,900 are accepted species names.

The Plant List contains 620 plant families and 18,167 plant genera.

The status of the 1,040,426 species names, are as follows:

<table>
<thead>
<tr>
<th>Status</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accepted</td>
<td>298,900</td>
</tr>
<tr>
<td>Synonym</td>
<td>477,601</td>
</tr>
<tr>
<td>Unresolved</td>
<td>263,925</td>
</tr>
</tbody>
</table>

Browse
Click on the major plant group of interest to explore the taxonomic hierarchy embedded within The Plant List.

- **Flowering plants**
- **Conifers, cycads and allies**
- **Ferns and fern allies**
- **Mosses and liverworts**

Work down the taxonomic hierarchy from **Major Group** (to find out which Families belong to each), to **Family** (to discover the Genera belonging to each) and finally **Genus** (to list the Species in each).
3) The Plant List

Collaboration: Kew, Missouri + many others
Alternative data sets merged by computer
Conflicting opinions detected and resolved where possible

Strengths:
- Complete a list as possible
- Prioritises data sets
- Browse feature
- Transparent & repeatable
- Record reliability indicated
- Links to other data

Weaknesses:
- Static – data 2012
- 25% names “unresolved”
- Quality compromised
- No geography
- No common names / images
- No services

www.theplantlist.org
Inter-relationships

World Checklist - IPNI - Tropicos

The Plant List - CoL
Top 5 things that Annoy Users

– Designed for botanists: Unintelligible to others
– Easy to misuse/ misinterpret
– Duplication: Don’t always agree!
– Unhelpful IT design, lack of options
– Name resources lack..
  • Common names, trade names, etc
  • Latin names as used in real world e.g.
    – Misspelt names in legislation
    – Names misapplied
Management issues

- Curation
- Long term support
- Audience engagement
- IT support and design
Curation

• Who curates?
• Significant Effort
  – IPNI 300K edits per year; World Checklist 100K
• Feedback- broaden the curation load
  – perhaps easier if community can be easily defined:
    – national or thematic user groups easier to engage with than unspecified broad global audience
Institutional support

• Needs long term security and time to build usage. Difficult with short term funding

• Strategic partnerships. Don’t try and do everything
Global Aggregation

- **Strengths**
  - Global
  - Comprehensive

- **Weaknesses**
  - Less responsive to users
  - Unfocussed
  - Harder to update

- Taxonomic Coverage
- Responsive
- Data Quality

USE
Audience Engagement

• Data isn’t everything (change in mind set)
• identify benefits to users,
• package data appropriately
• Think of the community as both users and suppliers
Medicinal Plant Names Services
MPNS Solution

Beta Release:

Please enter a NAME to search the MPNS resource:

Limit search to:

Fang ji  All names

8 records matched your search. These records relate to:

<table>
<thead>
<tr>
<th>Accepted scientific names</th>
<th>Scientific names as used in medicinal plant references</th>
<th>Non-scientific names</th>
<th>Medicinal plant references</th>
<th>All records</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>8</td>
</tr>
</tbody>
</table>

Accepted scientific names | Records referring to name
---------------------------|-----------------------------
Stephania tetrandra S.Moore | 5                           
Aristolochia fangchi Y.C.Wu ex L.D.Chow & S.M.Hwang | 3                           
Medicinal Plant Names Services

Each service is tailored to the needs of a specific audience and will enable more efficient and accurate use of plant names — reducing our clients’ costs and facilitating more effective communication and better targeted research across the health sector. MIPNS is actively engaged with partners in promoting global, industry-wide data standards.

To discover more about these services and how they can benefit your organisation, contact us at mpns@kew.org.

Validation
Checking and enriching your existing lists of plant names. We can correct spellings, propose updates to taxonomy, and enrich your dataset with all known synonyms for each plant. You will then be able to, for example, detect where you list a single plant under alternative names, or embed unique digital identifiers into your databases to facilitate maintenance of data and future updates.

Web Services
Connecting your IT systems to MIPNS electronically, thus enabling you to validate names as they are entered, refresh your databases as plant names change, build comprehensive names indexes and undertake intelligent data mining.

Vocabularies and Data Controls
Providing you with authoritative reference lists, ontologies, terminological controls and data sublists for use in enhancing your information systems locally.

Harmonisation
Enabling you to map your plant lists onto those plants cited in legislation or by other organisations or publications so as to understand overlaps, detect gaps and enhance communication.

Consultancies
We can provide 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. MIPNS also works closely with the Authentication Services for plant materials offered by Kew’s Innovation Unit (yellow link below).

Training
Specialist training courses for people working with medicinal plant names to enable safe and efficient working practices. These courses may be run at Kew or onsite for larger clients.
Going forward

• MPNS an examplar of way ahead: demonstrates relevance and importance of taxonomic research
• Outreach- complex and needs investment
• On going communication with users
• Relationship is two way, partners providing data – look for synergy and links to other resources
Conclusions

• Increased collaboration: technically, sociologically and in sharing resources
• Narrow the gap between providers and users, build mutual understanding- takes time, be patient.
• Data quality is vital, but we need to target resources to deliver benefits to intended audiences
• Be more business like: understand our costs and charge for premium targeted services