Communication Groups

1. *Invasive Species Specialist Group* (ISSG) of IUCN (The World Conservation Union) with newsletter *Aliens*. The ISSG in IUCN is attempting to establish a database of invasive species, though this database would be purely concerned with environmental weeds and would also include all biological invasives. Eventually IUCN proposes to link their databases to other similar databases, and it is possible that a network of databases may be created. The ISSG has also just produced its first newsletter (March 1995) *Aliens* which documents the effects of many biological invasions as well as providing a forum for discussion on this issue. *Contact*: Dr Mick Clout, IUCN Invasive Species Group, The Centre for Conservation Biology, University of Auckland, Tanaki, Private Bag 92019, Auckland, New Zealand. Tel: +64-9-3737599 Fax: +64-9-3737042
E-mail: m.clout@auckland.ac.nz

2. *Exotic Pest Plant Councils* (USA). *Contact*: Dr Faith Campbell, (address unknown) USA.
E-mail: +1-202-682-1331

3. *California Exotic Pest Plant Council* (CalEPPC) - USA. *Contact*: Dr Steve Harris, USA.
E-mail: sharris@igc.apc.org

4. *Weeds of the World*. *Contact*: Dr Philip Bacon, Department of Plant Sciences, South Parks Road, University of Oxford, Oxford, OX1 3RB, UK. Tel: +44-1865-275066
Fax: +44-1865-275146 E-mail: PBacon@vax.ox.ac.uk

Other Relevant Groups

1. *Taxonomic Database Working Groups* (TDWG). TDWG has published standards on data exchange, world geography and names of taxa; has endorsed standards on authors, bibliographic citations, herbarium code designations, phytogeographic units and economic use descriptions (Bisby, 1994). *Contact*: Dr Frank Bisby, Biodiversity and Bioinformatics Research Group, Department of Biology, University of Southampton, Southampton, SO16 7PX, UK.
Tel: +44-1703-595000 Fax: +44-1703-592444 E-mail: F.A.Bisby@southampton.ac.uk

2. *National Weeds Strategy* (Australia). *Contact*: Dr Jim Cullen, (address unknown) Australia.
E-mail: jimc@ento.csiro.au

3. *International Organisation for Plant Information* (IOPI). *Contact*: Dr Richard Pankhurst, Royal Botanic Gardens, Inverleith Row, Edinburgh, EH3 5LR, Scotland, UK.

4. *Office of Technology Assessment* (USA). *Contact*: OTA (US Congress Office of Technology Assessment), 600 Pennsilvannia Av WSE, Washington DC, 20510-8025, USA.
Tel: +1-20510-8025

5. *TIGER* (*Terrestrial Initiative in Global Environmental Research*). The Working Group 4 of TIGER has within its list of high-priority research areas: "collating and manipulating datasets of the distribution of species and physical attributes of the environment to define and predict impacts of global change at landscape and regional scales". *Contact*: Dr Clive Cummins, Institute of Terrestrial Ecology, Monks Wood, Huntingdon, Cambridgeshire, PE17 2LS, UK.
Tel: +44-1487-773381

6. *AETFAT - Association for the Taxonomic Study of the Flora of Tropical Africa.* A list of members is being stored in the database. *Contact*: Prof dr LJG van der Maesen, Department of Plant Taxonomy, Herbarium Vadense, Agricultural University, P.O.Box 8010/6700 ED, Wageningen, Netherlands. Fax: +31-8370-84761 E-mail: Jos.vanderMaesen:algem.pt.wau.nl

RESULTS

Needs assessment study

In the initial project document one of the main project outputs was highlighted as being a needs assessment study for an invasive plants database. Opinions on database format were sought from a number of scientists working in related research in an attempt to assess the perceived requirements of any database.

The overwhelming response from researchers in every field was that the current knowledge of the status of *weeds* needed to be brought together in some type of information system so that data could be shared by all interested parties. The reasons given for this were many, though they fell broadly into the following categories:

1. To assist scientists in collaborative research on specific plant species.
2. To be able to record and monitor the present distribution of known *weeds*.
3. To attempt to predict the spread of invasive plant species.
4. To guide quarantine control in regulations for transportation of plant materials.
5. To collate data that can used to demonstrate the environmental and agricultural impact of weed species and so raise awareness of the problem.

Storing the data on all categories of *weeds* in one database did not meet with the approval of all those consulted. The reasons given for this were:

1. The disparity between the well documented agricultural weeds and the comparatively poor documentation of the environmental weeds.
2. Invasive plants were considered not to be weeds, but potential weeds.
3. Agricultural weeds are predominantly herbaceous, whereas environmental weeds are mostly woody.
4. Wide variations in disturbance regimes between agricultural and environmental weeds.
5. The differences between agricultural weed scientists and environmental weed researchers and the origin of funds for research.

However, it was generally felt that a database of all *weeds* was accomplishable and there was a need for all researchers to collaborate to a far greater extent. This is particularly important since there is currently a shift of emphasis away from funding traditional weed science research to environmental weeds, due to public awareness. A greater effort is required to encourage taxonomists, agronomists, ecologists, botanists and conservationists to communicate.

Suggestions for database design were also numerous, owing to the diverse set of database requirements; each researcher having their own set of data which they wanted to see included.

Naturally it is almost impossible to incorporate all of these data fields into one database. The priorities for database content and design, from the views of those consulted, were:

1. To ensure taxonomies are correct, with proper inclusion of synonymy.
2. To record distribution on a regional or agro-ecological zone basis.
3. To endeavour to gauge weediness/invasiveness.
4. To outline potential control methods.
5. To provide a comprehensive bibliography of all relevant research.
6. To be available in hard copy as well as on-line or on disk.
7. To include GIS linkage; vital for future monitoring.

A number of possible problems were put forward which might hinder the construction of this type of database. Not least of these was the fact that it is extremely difficult to classify plants according to their status or invasiveness, because this largely depends on habitat. At present the only good predictor of invasiveness is the behaviour of a plant outside of natural habitat.

Consequently, the perceived options for database design fell into four categories:

1. A compromise database, to encompass as many data fields as practical within a single entry (which might in the end please no one).
2. A large database specifically designed for invasive weeds (though no one believed that this type of database could be achieved, due to a number of reasons).
3. A very basic and simple database with the minimum number of data fields. If a database is to be created for *weeds* worldwide then a simple database is required with additional information sources included, i.e. on research. If a database is kept simple then more people are likely to use it, and it can be added to or adapted.
4. A network of existing databases, which might be an on-line centrally managed information system of "metadata" on regional databases, which should all be of a similar format.

The database design that found most favour was option 4. It was possible that this was because each country has its own methodology for dealing with invasive plants; its priorities are different; and depend on a wide set of variables. It may be easier to encourage collaboration if there was a centralised information system which would allow people to communicate and use the databases that are already exist, or are being constructed. A new database should not be created unless its end use justifies it, and it was felt that there is likely to be more regional use of a database than there would be global applications.

It was generally believed that long-term databases will not receive funding because they are costly to produce and maintain. The survey showed that current individual database initiatives are important but need linking, and not amalgamating. It is possible that organisations may be more willing to contribute funds to a centralised information system.

Respondants favoured the establishment of an independant committee on invasive species, or one formed within a present organisation. It was suggested that the International Weed Science Society may be the most suitable organisation to take on this role.

Database Content

Because the range of potential end-users of any *weed* database is extremely wide, then the requirements of that database will also be very large, if all the users are to be satisfied. Whilst in theory a database incorporating all possible data fields would be preferred, in practice this is not possible. A compromise needs to be found.

Whilst it is not possible to attempt to recommend the definitive database format in this paper, there will be some data sets that are common requirements to all the potential database users, and beyond this further data sets can be listed according to perceived importance.
Even in its simplest form a database for *weeds* needs to be a relational database with linked individual modules, enabling searches by taxonomic classification, synonyms, distribution, and habitat classification.

For some of the data fields (e.g. those within habitat, plant form and distribution) a coding system would be required to minimise the text and so reduce the required space. However, to ensure that the database remains "user friendly" and mistakes are minimised during data entry, actual names, or descriptors would need to be used in the majority of instances. The recommended minimum amount of data content required in a purposefully designed modular, relational, *weed* database should contain the following:

Module	Data field/data type
1 Family:	Family
2 Genera:	Genus, Species, Authority.
3 Synonymy:	Synonyms
4 Distribution:	Origin, Abundance, Agro-ecological zone, Country, Region.
5 Morphology:	Life form, Descriptors (flower, leaf, fruit, seed, roots).
6 Habitat:	Climatic factors, Edaphic factors, Disturbance factors.
7 Character (plant):	Lifespan, Uses (human), Light requirement, Reproduction.
8 Disturbance:	Type, Level, Frequency.
9 Control:	Type, Success rate.
10 Information:	References, Abstracts, Research Organisations, Individual contacts.

Naturally, there is great scope for increasing the datafields and dividing the modules further. The actual database software or computer language used for database construction is important. However, as the database market place is constantly changing it is impractical to make a specific recommendation. Despite this, guidelines for any database selected are possible; it would require a) to have collapsible data fields, b) to be IBM compatible, c) to run under *Windows*, and d) to be user friendly.

If a standard commercial database package needs to be distributed to potential database users at a later date then the software selected should be purchased as a run-time module. This means that the software can be distributed to the participants at the same time as the stored data, without them having to purchase the software themselves.

Holm *et al.* Database (World Weeds Database)

The World Weeds Database (WWD) prepared at OFI (Oxford Forestry Institute) is based on the information contained in the book "A Geographical Atlas of World Weeds" by Holm *et al.* (1979).

A program was developed to retrieve valuable information from this book and make it available to any person who might be interested in knowing the distribution of the most common and worst weeds. The book is a comprehensive list of weed species of the world and their distribution and even though it does not include all the species for any geographical region, it is a good information source for field based scientists.

The WWD displays information on over 2,400 weed species around the world according to their rank of importance. At present the WWD represents only part of the book (up to 'G' in the family list), though attempts are being made to recover the remaining data. The ranking includes five different categories each one identified with a number as follows:

5 - Serious weed
4 - Principal weed
3 - Common weed
2 - Present as a weed (the species is present and behaves as a weed, but its rank of importance is unknown)
1 - Flora (the species is known to be present in the flora of the country, but confirming evidence is needed that the plant behaves as a weed)

The WWD offers not only information on each individual species (using this rank of importance and its distribution around the world), but also provides the enquirer with four different ways of looking at the information. Database searches can be made as follows:

1. Country: The country search form provides a list of 123 countries for the user to choose from.
2. Plant: This form provides a full species list with over 2400 records.
3. Genus: The user can select a genus name from a list of 270 different genera.
4. Family: There are 140 family names.

The information retrieved from the book is stored as a relational database in different tables in *Access 2.0*, as a relational database. This stores and retrieves information according to the relationships previously defined and allows error checking.

Weeds of the World Project can be viewed on the home page of Oxford Forestry Institute web server and located at the URL: http://ifs.plants.ox.ac.uk/wwd/wwd.htm

Future database/information system

The options are as follows:

1. *Creation of a new database*. As a large number of similar botanical databases already exist, or are being compiled, it seems impractical to consider creating a totally new database specifically for

weeds and/or invasive plants. It is also likely to be very difficult to obtain funding for such a project. With this in mind, only the options of updating and modify old databases, or adding to existing databases, have been considered.

2. *Adapt existing Holm et al. database*. Given that it has been shown to be possible to adapt the Holm *et al.* database to a desired format, the option of going ahead and making the further modifications and putting the database "on-line" seems an attractive one. It would still be necessary to add the remaining data prior to up-dating the entries for each species. This would depend upon the publishers consent, and being able to obtain the remaining material from the original database.

If this database is developed further, the site at which it is based is not important for the actual work, provided that the required facilities and expertise exist. However, its situation may be important when considering the desired involvement of a wide range of researchers. To locate an invasive species database at a site which has traditionally been associated with agricultural weed science research may not appeal to botanists, ecologists or conservationists, consequently their collaboration may be lost. An organisation which has an international reputation and has had, or is moving towards, involvement with both agricultural and environmental weeds would be the optimum choice. However, because more complete and purposefully constructed databases exist it may be preferable to import the data from the WWD into such a database, once all the data has been retrieved.

3. *Create a weeds module in a larger external database*. An existing database, or one in the process of being created, could possibly be adapted to include a module on *weeds*. Some of the databases that have that potential were listed previously, though it is likely that other suitable opportunities may arise for collaboration. A module created in this way would obviously be written and structured in a similar format to its parent database and would be compatible with like modules and similar databases. The associations of any such module with its parent database would be far more important than where it was located and so it could be maintained at any site.

4. *Establish a "metadatabase" of existing databases*. Many scientists in the field of *weed* research believe that a research network and information system, designed to enable them to exchange data and access other research more easily, would be more beneficial to their research than if a new database was created..

Several databases of relevance to *weed* research exist already. However, researchers are often ignorant about their existence, or are unable to access the information they require. These databases could be linked via a metadatabase; i.e. a central database of databases (metadatabase) could be created. It would function as an information system, linking all the weed researchers through a central network, which could be used by anyone seeking information. Through the existence of such a network and the resulting communication, confidence in the network within the research community could be established.

For this type of system donors for funding would have to be sought initially to finance the creation of the network. However, once established it could be possible to charge a fee, to cover costs, if this was deemed appropriate. Subscription rate could be varied according to the status of the organisation, in a similar fashion to *BG Recorder* or *PRECIS*.

5. *Modify an existing database of an appropriate organisation.* There are a number of botanical databases that might lend themselves for adaption to include *weeds*, or exist in an altered form specifically for invasive plants. Although this would be dependant on the organisation responsible for holding the original database, it would be relatively straight forward to use many of the existing data fields and convert others so as to facilitate the entry of data required in a *weeds* database (e.g. *BG-Base* or *BG-Recorder* for botanical data or the *Bushweed 2 Database of Environmental Weeds of Australia which is designed for recording data on weed* species).

6. *A combination of options and database dissemination.* A combination of more than one of the above alternatives could be considered. For instance, the creation of a metadatabase could easily be put into action at the same time as alternatives 2 or 3. If this combination was employed then a more concrete database could evolve to become the focal point for *weed* data which would meet both long and short term goals.

It is essential that in future all data is stored in electronic form, in a widely used database format to enable it to be incorporated into other associated databases. This will not only aid database construction and development, but assist in improving the dissemination of data. All modern databases should work towards being made available on-line and over the internet. In some instances, especially where data is relevant to development issues in poorer nations, a hard copy of the database would be essential. Most often this requires printing the data, although, increasingly CD-Roms are becoming available in centres of development.

Information needs to be disseminated to the end user in the cheapest way. Unlike agricultural weeds within farming systems, the cost of environment weeds to natural ecosystems cannot be measured accurately. This point needs to be taken into consideration when any commercial venture involving *weed* data is assessed. All information should be easily available; so that research scientists have the fewest possible barriers to collaborative research.

CONCLUSIONS

Research scientists, and other interested organisations, urgently require some type of system to access data on weeds and invasive plant species. Improved communications and information availability are the two immediate priorities. Data needs to be stored in a form in which it is easily retrievable by all those who need to access it.

The establishment of a communication network, linking institutions and individuals concerned with *weed* research, would meet these requirements. This network has to be administered by an appropriate organisation, with experience in plant databases, though the site at which it is based is not important, providing that it has the necessary computer equipment and is on-line within the internet system. The main objective of such a network would be to facilitate communication between relevant institutions and individuals. However, in addition to this task, the networking organisation, or another organisation collaborating with the network organisation, should also set out to achieve the following objectives: a) facilitate the communication of relevant institutions and individuals; b) assist in the transfer of plant data; c) raise public awareness; d) recommend a common database format for further database construction; e) work towards constructing, or adapting, a metadatabase.

Any networking of *weed* databases has to be undertaken by an organisation which is respected within this sphere of research and preferably has experience in plant database management. The *International Weed Science Society* (IWSS) is an obvious choice, though it has not previously taken on such role. So far there have been only a handful of organisations (governmental or NGO) that have made substantial progress in coordinating research on weeds and invasive plants.

The only international organisation to have implemented policies on weeds or invasive plants is the IUCN who have identified the problem of biological invasions and have also established an *Invasive Species Specialist Group* (ISSG) that has recently started a newsletter ("Aliens"). In addition to this it is presently trying to create a database of invasive species. IUCN also has experience with database creation. Whilst not all of the IUCN's work on invasive species meet the objectives of the database proposed in this project there is a great deal of common ground which suggests that IUCN may be well placed to take on the database network. The in volvement of the International Weed Science Society (IWSS), though not a funding body, would enhance the standing of a *weeds* database, thereby improving the chances of success in obtaining funding. Naturally there are other organisations that could also be viewed as potential collaborators, such as BGCI which currently exchanges data with IUCN.

It is vital that available data on all categories of weeds and invasive plants is made accessible to everyone who requires it. This means that databases such as the WWD need to be developed further so that all the data is retrieved and made fully available in electronic form. Either it can be made available in its own right (either on-line or on CD-Rom), or used to assist members of a *weeds* network, through being incorporated into a purposefully constructed database. Recommendations for database format should be made in conjunction with other database working groups such as TDWG. Ultimately, all plant databases are of potential interest to everyone from *weed* researchers to plant quarantine officials, so it is important that the appropriate standards and database compatibility are established at the earliest opportunity.

To achieve the most efficient use of time and resources an existing botanical database (such as BG-Base or BG Recorder) should be adapted for storing data on *weeds*. However, the choice of database should be that of the networking organisation, in collaboration with other organisations.

REFERENCES

Bisby, F (1994) Global Master Species Databases and Biodiversity. *Biology International*. July 1994: 33-40.

Holm, L R; Pancho, J V; Herberger, J P; Plucknett, D L (1979) *A Geographical Atlas of World Weeds*. Wiley and Sons, London.

Smith, C (1992) Impact of Alien Plants on Hawai'i's Native Biota. In: *Hawai'i's Terrestrial Ecosystems: Preservation and Management*. C P Stone; J M Scott (Eds) University of Hawaii Press, Honolulu.

Swarbrick, J T; Skarrett, D B (1994) *The Bushweed 2 Database of Environmental Weeds in Australia*. Gatton College, The University of Queensland, Queensland, Australia.