# PLANT OCCURRENCE AND STATUS SCHEME

## *A Standard for Recording the Relationship between a Plant and a Place*

**compiled by the**
[World Conservation Monitoring Centre](#)

An [International Working Group on Taxonomic Databases](#) Standard

Biodiversity
Information
Standards
TDWG

---

## *Table of contents*

## Acknowledgements

## Introduction

<u>Standards</u>

The development of agreed standards for the management of information is a key issue in ensuring that data can be shared on as wide a level as possible, and as easily as possible. This is particularly important for data maintained in computerised format on databases, if this data is to be made available to many users.

Considerable debate may be involved in the preparation of agreed standards which, at first sight, may appear deceptively simple. However, all those involved in management of data will be well aware of the problems that exist if there is no clear, agreed standard to follow for a particular issue. In the worst instances, lack of agreed standards can mean data sets are managed in such a way that wasteful re-keying in of data is needed.

## POSS

The current standard, the Plant Occurrence and Status Scheme (POSS) has been developed to provide standard terms for recording the occurrence of a plant taxon for a specific location.

As with all standards, POSS aims to provide both a clear, un-ambiguous framework for information management, and at the same time provides a mechanism for information exchange.

It is worth noting, that although the standard has been developed with the management of plant information in mind, the standard is equally suitable for the management of animal information, although animal movement and migration are not provided for.

## Development of POSS

Many possibilities arise in the design of a scheme such as POSS. The current scheme results from discussions between many WCMC staff over many years, during which time it was recognised that no one scheme was totally defensible, totally comprehensive or totally objective. However, many fields of the current scheme have already been used extensively at the World Conservation Monitoring Centre over the last 5 years, for the management of data on plants threatened at the national and global scale, thereby providing a good test for POSS as it now stands.

POSS has been developed over a lengthy period. Earlier drafts, circulated under the title "Plant Existence Categorization Scheme (PECS)", also covered "Nature of Cultivation", "Re-establishment" and "Degree of Threat" and suggested covering "Frequency" and "Distribution Patterns". These have since been considered inappropriate for inclusion in POSS since they go beyond the scope of the current scheme, and are too specialized.

The field "Endemism", presented in previous drafts, has been omitted from this version. The concept of endemism, however, is still an integral and important part of POSS, see the field "Distribution Completeness Information".

A preliminary version 1.0 of POSS was prepared in draft format in 1989, but was not published.

As with all standards, development of further versions of this standard may be appropriate, following use by a wide range of users. Comments and feedback for future versions are welcome. Please send all comments to the TDWG secretariat, *via* the TDWG representative at the World Conservation Monitoring Centre, 219 Huntingdon Rd, Cambridge CB3 0DL, UK.

Relationship to other data standards

Any standard for data management needs to be designed to take into account other existing related standards. Elements within POSS naturally relate particularly to elements of other TDWG agreed standards.

The "Degree of Threat" is excluded from POSS, because it is covered elsewhere. The International Transfer Format for Botanic Garden Records TDWG Standard 1 (ITF), covered conservation categories as first defined by the World Conservation Union IUCN. These categories have recently been the subject of much debate, and have now been superseded by publication of IUCN's revised threat categories *IUCN Red List Categories* (IUCN, 1994), prepared by the IUCN Species Survival Commission.

The ITF is, in turn, currently being revised, and in the draft Version 2.0, provision is made for these new IUCN categories. The draft Version 2.0 also makes provision for the inclusion of fields determined by POSS.

General Principles of Use

The Plant Occurrence and Status Scheme (POSS) provides the link between a plant taxon and any geographical area. The size of the geographical area is un-important. The scheme provides fields to record whether the taxon in question is present or absent from a particular area, and its status within that area. Status refers to whether the taxon in question is native, introduced or cultivated or in the area under consideration. Additionally, POSS records how complete the information set is both for the local area, and on the global scale.

NB The term "status" as used in this document refers to the plant's origin and should not be confused with the quite separate issue of THREAT status. The standard published by IUCN in 1994 provides details of IUCN's updated classification of threat status, developed over a period of several years. This

scheme is designed to replace the original less objective classification scheme also produced by IUCN.

General principles for following POSS:

■ The taxon should be a known and taxonomically distinct entity, although this does not preclude maintenance of occurrence data on entities that are nevertheless still taxonomically debatable. Normally POSS will be applied to species and infra-specific taxa; however, it may be applied to any taxon, for example a family.

■ The area must be defined, i.e. it must have definite boundaries, but it can be of any size. The standard should work just as well for a small nature reserve as for a continent, although probably it will be most useful for data on the occurrence of plants in countries, states and islands. POSS is designed so that areas can be treated hierarchically without upsetting the standard.

The data managed using this standard have been divided into their constituent parts in line with good database design. For the same reasons, all data fields and values within those data fields are non-overlapping. Wherever values within or between data fields are mutually exclusive, there is usually an indication in the text. Within any data field, there should be a value to represent every circumstance. The seven data fields forming the current standard, are described in the Appendix.

Mandatory fields

Of the seven data fields, only Data Field 1 **Occurrence** is mandatory. The others are optional. For example, a database on plants in a national park may not be considered an appropriate place in which to record the cultivated status of a particular plant. In this example, Data Field 4 **Cultivation**" would be omitted.

Some fields are mutually dependent. If Data Field 4 **Introduced Agency** is included, then Data Field 3 **Introduced Status** must also be included.

Omitted fields

It is important to note that if any data fields are considered inappropriate for use in a particular database, they may of course be omitted. However, it is then not acceptable to merge data relevant to an omitted field, under other headings. For example, if the Data Field 5 **Cultivated Status** is omitted, it is not acceptable to redefine the Data Field 3 **Introduced Status** to include both truly introduced and cultivated plant records.

Omitted values for a field

Many database designers may not wish to use all the values in each of the data fields they have chosen. For example, many will feel it unwise to permit the use of the value **ASSUMED PRESENT** in Data Field 1 **Occurrence**. This is quite acceptable. It is, however, equally important that, as with omitted fields, data for non-implemented values are not merged under implemented ones. For example, if in Data Field 1: **Occurrence**, the value for **RECORDED AS PRESENT IN ERROR** is not implemented, it is not permissible to utilize the value for **DOUBT ABOUT PRESENCE** to include this concept, as this would corrupt the usage of **DOUBT ABOUT PRESENCE**.

NB Experience at WCMC demonstrates the benefit of using the **RECORDED IN ERROR** value. Deletion of an erroneous record gives rise to two problems: Firstly, the original data may be inadvertently re-entered. Secondly, deletion of a record in an indexed relational database is liable to create problems, with cross-referenced files holding meaningless records linking back to a now non-existent record in the initial file.

Additional fields or values

If it is found necessary to define extra values or data fields, database managers are asked to contact WCMC so these values or data fields may be considered for future versions of the standard. For the data fields concerning status information, there is a value **NONE OF THE ABOVE**. If it is found that this value has to be used often, then either the remaining values are being misinterpreted, or this is an indication that one or more new values may be needed.

Transfer Values

Many databases may wish to use a different internal format for managing the data fields.

For example:

■ transfer value "1 " may be used instead of "N" for the value **NATIVE** in Data Field 2: **Native Status**

■ transfer value "Y" meaning "yes" for **PRESENT** in Data Field 1: **Occurrence**.

Frequently, an internal format may be adopted to suit the language of the user - providing a memorable mnemonic character. For example, in French "I" for Indigène, would be more appropriate for a user, rather than "N" for Native).

The important elements in the standard are not the codes chosen in each case but the definition and delimitation of the values themselves. As in all cases of substitution of codes for values in databases, it is the concepts behind the values that matter, not the codes.

The standard remains intact whatever set of codes are employed to represent the values expressed in this document, providing:

■ the codes are unique;

■ the values they represent are intact; and

■ the codes are used consistently throughout the entire database.

*A Summary of POSS*

The POSS categories form the necessary link between:

■ Any named plant taxon; and

■ Any defined area

To do this, seven **Data Fields** are defined:

OCCURRENCE INFORMATION

| Data Field 1: **Occurrence** | Does the plant occur in the area ?(Regardless of whether the plant is Native, Introduced or Cultivated in the area.) |
| --- | --- |

STATUS INFORMATION

| Data Field 2: **Native Status** | Is the plant native in the area? |
| --- | --- |
| Data Field 3: **Introduced Status** | Is the plant introduced in the area ? |
| Data Field 4: **Introduction Agency** | Is the introduction considered natural? |
| Data Field 5: **Cultivated Sta tus** | Is the plant cultivated in the area? |

DISTRIBUTION COMPLETENESS INFORMATION

| Data Field 6: **Area Distribution Completeness** | Do the area records in the database represent the distribution of the plant within the compass of this area record? |
| --- | --- |
| Data Field 7: **World Distribution Completeness** | Do the area records in the database represent the complete world distribution of the plant? |

NB Data Fields 1-6 are used for each plant-area occurrence Data Field 7 is used only once for the taxon

## References

Bisby, F. 1994 *Plant Names in Botanical Databases. Plant Taxonomic Database Standards No.3* Version 1.00, December 1994. 30pp. Hunt Institute for Botanical Documentation, Pittsburgh.

Brummitt, R.K. 1992. *Vascular Plant Families and Genera*. 804pp. Royal Botanic Gardens, Kew.

Brummitt, R.K. & Powell, 1992. *Authors of Plant Names*. pp. Royal Botanic Gardens, Kew.

Hollis, S. & Brummitt, R.K. 1992. *World Geographical Scheme for Recording Plant Distributions, Plant Taxonomic Database Standards No. 2*. 105pp. Hunt Institute for Botanical Documentation, Pittsburgh.

IUCN/WWF 1992. *The International Transfer Format for Botanic Garden Plant Records Plant Taxonomic Database Standards No. 1*. 70pp. Hunt Institute for Botanical Documentation, Pittsburgh.

IUCN. 1994. *IUCN Red List Categories.* 21pp. IUCN, Gland, Switzerland.

## Data Fieds

### *OCCURRENCE INFORMATION*

### Data Field 1: Occurrence

This data field defines whether or not the plant occurs in the area concerned, regardless of its status.

| Value | Suggested Transfer Value |
|---|---|
| **PRESENT** | P |
| **ASSUMED PRESENT** | S |
| **DOUBT ABOUT PRESENCE** | D |
| **EXTINCT** | E |
| **RECORDED AS PRESENT IN ERROR** | F |

| ABSENT | A |
|---|---|
| NO INFORMATION | - |

*General*

This data field refers to the status of the taxon at the present time, not at the time at which the records were made. It is thus technically an assumption by the author of present status based on the records available at the time the plant-area record was made in the database. (This is exactly comparable to Floras and Checklists, which purport to outline which plants occur in their territory, rather than to say which plants occurred at the times of individual records.) Therefore, if the database is updated only very seldom, it may be useful to associate a date alongside the time the plant-area record was made or approved, although this is not required for the purposes of this standard.

The status of the plant, whether native, introduced or cultivated, is irrelevant to this data field.

The standard provides in detail for recording doubt about a plant's presence or status. It is to cover the various situations of doubt that the **Occurrence** data field is mandatory: for example, if a plant is definitely present in the area, but its status is unknown, the Data Field **Occurrence** would be set to **PRESENT** and the remaining three data fields set to either **DOUBTFULLY NATIVE/DOUBTFULLY INTRODUCED/DOUBTFULLY CULTIVATED** or **NO INFORMATION**, or combinations of these.

*Values Explained*

**PRESENT:** The plant is present in the area, as authenticated in the botanical literature or by field observations or by herbarium specimens or by any combination of these. One or more of these records must have been made in recent years to justify the assumption of present occurrence. Consequently, the author accepts that the taxon is still there until there is evidence to the contrary.

**ASSUMED PRESENT:** It is highly probable that the plant does occur in the area, but there are no definite records for its presence there. An example of this situation is where positive records indicate that a taxon is widespread in all tropical African countries except one, where perhaps no Flora is available, and so the plant is, therefore, assumed to be present in that one country too.

**This value is very dependent on interpretation. Floras and Checklists rarely mention species assumed to occur, though doubtful records are usually included. For these reasons, most databases will wish to exclude using this value, but it is included here for the sake of completeness.**

N.B. ASSUMED PRESENT should not be confused with the concept of expected future occurrence: the latter is where a plant does not occur in a certain area at present but it is likely to spread into that area in the future. This concept is not covered in POSS, which refers only to taxon-area records for the present time.

**DOUBT OVER PRESENCE:** There is doubt as to whether the plant presently occurs in the area or not because there are records for the plant in the area but they are all in doubt. Certainty about any one record would, of course, indicate that the plant is present.

**DOUBT OVER PRESENCE** can, for example, include situations where:

> ■ All records are very old, perhaps over 100 years old or more, and the vegetation of the area concerned has changed greatly since those records were made.
>
> ■ All relevant herbarium specimens on which occurrence was based are of uncertain identity (eg. lacking flowers or fruits).
>
> ■ It is uncertain whether the specimens were collected within the area or not, either due to uncertain localisation, such as the result of the inability to interpret an author's use of geographical areas, or due to boundary changes since the record was made.
>
> ■ A lack of repeated searches of known and likely areas necessary to demonstrate that the plant is either present (extant) or presumed extinct in the area.

This value does **not** refer to a certain record (within the area defined) whose precise locality is uncertain.

**EXTINCT**: The plant was once in the area (previously **PRESENT** or **ASSUMED PRESENT**) or may have once been in the area (previously **DOUBTFULLY PRESENT**) but is now known to be extinct there, or extirpated. The criterion for extinction is that the plant was not found after repeated searches of known and likely areas (i.e. sites within the area covered by the record), even though the plant may be extant elsewhere. The concepts of extinction and extirpation are therefore treated identically within POSS.

**RECORDED AS PRESENT IN ERROR**: The plant has been recorded as present in the area concerned but all of those records have been disproved or discounted.

**ABSENT**: There is no reason to believe the plant has ever occurred in the area concerned. No records exist for the plant's occurrence in the area, neither can it reasonably be "assumed" to occur there. Evidence should exist, in some form, to indicate that the plant does not occur in the area concerned.

**NO INFORMATION**: There is no information as to whether the plant occurs in the area concerned .

## *STATUS INFORMATION*

### Data Field 2: Native Status

Whether or not the plant is Native in the area concerned.

| Value | Suggested Transfer Value |
|---|---|
| **NATIVE** | N |
| **ASSUMED** TO BE NATIVE | S |
| **DOUBTFULLY** NATIVE | D |
| **FORMERLY** NATIVE (EXTINCT) | E |
| **NOT** NATIVE | A |
| **RECORDED AS** NATIVE IN ERROR | F |
| **NO INFORMATION** | - |
| **NONE OF THE ABOVE** | U |
| **NOT APPLICABLE** | X |

*General*

This data field refers to whether the taxon is indigenous within the area concerned, or whether there is doubt over native status, or if the taxon is known definitely not to be native.

*Values Explained*

**NATIVE:** The taxon is indigenous within the area concerned, and there is evidence that it has a history there, or circumstantial evidence of occurrence in one or more natural or semi-natural plant communities.

**ASSUMED TO BE NATIVE:** Assumed to be Native in the area concerned. The assumption may be made, for example, on the basis of evidence of vegetation community structure etc.

**DOUBTFULLY NATIVE:** There is doubt as to whether the status of the plant in the area concerned is native, as defined above, or not. All records about the native status of the plant in the area are in doubt.

**FORMERLY NATIVE (EXTINCT):**The plant is native, doubtfully native or assumed to be native in the area concerned and has become Extinct as such. The criterion of extinction is that the plant was not found (as a native) after repeated searches of known or likely areas (i.e. sites within the area covered by the record), even though the plant may be extant elsewhere.

 NB The issue of recording extinction of a taxon is also dealt with under separate standards, notably the standard produced by IUCN (1994) *IUCN Red List Categories*.

**NOT NATIVE:** The plant is definitely not native (as defined above) in the area concerned.

**RECORDED AS NATIVE IN ERROR:** The plant has been recorded as native in the area concerned but all of those records have been disproved or discounted. A known fallacious native record must have been made, and it must be known that the plant does not occur as a native in the area to use this record.

**NO INFORMATION:** There is no information as to whether the plant is native or not in the area concerned.

**NONE OF THE ABOVE:** This is included so as to identify any situations that may arise which cannot be covered by the above values.

**NOT APPLICABLE:** This data field does not in any way apply to the plant in this area because Data Field 1: Occurrence has the value ABSENT.

Data Field 3: Introduced Status

Whether or not the plant is introduced in the area concerned, irrespective of whether the introduction was known to be caused by man, either deliberately or accidentally, or by any natural means.

| Value Suggested | Transfer Value |
| --- | --- |
| **INTRODUCED** | I |
| **ASSUMED TO BE INTRODUCED** | S |

| | |
|---|---|
| **DOUBTFULLY INTRODUCED** | D |
| **FORMERLY INTRODUCED (EXTINCT)** | E |
| **NOT INTRODUCED** | A |
| **RECORDED AS INTRODUCED IN ERROR** | F |
| **NO INFORMATION** | - |
| **NONE OF THE ABOVE** | U |
| **NOT APPLICABLE** | X |

NB If a plant is introduced, then an entry **must** be made in this data field.

*General*

This data field refers to whether the taxon is introduced within the area concerned, or whether there is doubt over introduced status, or if the taxon is known definitely not to be introduced.

*Values Explained*

**INTRODUCED:** The plant has been recorded growing in an area that is outside of its assumed true and normal distribution. This implies evidence that the plant did not formerly occur in the area, but has "recently" arrived. The definition of recently is deliberately vague, to accommodate as many circumstances as possible. The plant must not be in cultivation, but may occupy any habitat within that area, whether natural, semi-natural or ruderal. This term does not mean (or include) "Introduced to Cultivation". The means of introduction, whether by man or any natural means is irrelevant and may be unknown.

**ASSUMED TO BE INTRODUCED:** There is doubt as to whether the Status of the plant in the area concerned is Introduced, as defined above, or not. All records about the introduced status of the plant in the area are in doubt.

**DOUBTFULLY INTRODUCED:** There is doubt as to whether the Status of the plant in the area concerned is Introduced, as defined above, or not. All records about the introduced status of the plant in the area are in doubt.

**FORMERLY INTRODUCED (EXTINCT):** The plant is introduced, doubtfully introduced or Assumed to be introduced in the area concerned and has become Extinct as such. The criterion of extinction is that the plant was not found (as an Introduction) after repeated searches of known and likely areas (i.e. sites within the area covered by the record), even though the plant may be extant elsewhere.

**NOT INTRODUCED:** The plant is definitely not introduced (as defined above) in the area concerned .

**RECORDED AS INTRODUCED IN ERROR:** The plant has been recorded as introduced in the area concerned but all of those records have been disproved or discounted. A known fallacious introduced record must have been made, and it must be known that the plant does not occur as an introduction in the area.

**NO INFORMATION:** There is no information as to whether the plant is Introduced or not in the area concerned.

**NONE OF THE ABOVE:** This is included so as to identify any situations that my arise which cannot be covered by the above values.

**NOT APPLICABLE:** This data field does not in any way apply to the plant in this area because Data Field 1: Occurrence, has the value **ABSENT.**

Data Field 4: Introduction Agency

By what means did the introduced plant arrive in the area concerned?

| Value Suggested | Transfer Value |
|---|---|
| **INTRODUCED BY HUMANS** | M |
| **INTRODUCED BY NATURAL MEANS** | N |
| **NO INFORMATION** | - |
| **NOT APPLICABLE** | X |

*General*

This data field is optional, even if Data Field 3: **Introduced Status** is implemented. It serves to further qualify introduced species, by allowing information on the nature of their introduction to be managed. However, if this data field (**Introduction Agency**) is used, then Data Field 3: **Introduced Status** is mandatory.

It is intended to deal with the situation of taxa arriving in areas in which they were previously unrecorded or assumed to be present, by clearly natural means, such as wind or current dispersal. Inevitably the use this data field is subjective, since all native species distributions as we know them today must be "introductions" by natural means. The use of this data field must therefore be restricted to cases where there are clear records showing the recent arrival of a species in an area

in which it has never been recorded before, along with evidence about the means of arrival.

The use of data field 4, with the value **INTRODUCED**, and data field 5 with the value **INTRODUCED BY NATURAL MEANS** provides an interpretation of an "alien" species concept used by some botanists. Since this term is used particularly inconsistently, its definition has been deliberately avoided in POSS.

It is recognized that the distinction between indirect introduction by man, and natural introductions will often be a difficult one to make. for example, introduction by animals when the animal's own distribution has been affected by man could be construed as an indirect introduction by man, or a nature introduction. In each case of this sort it will be necessary to examine the evidence to ascertain the degree of influence of man and its relevance.

*Values Explained*

**INTRODUCED BY MAN:** The plant is introduced and the agent is known to be man, whether accidental or deliberate. This example includes, for example, taxa that have escaped from cultivation and those that occur as a result of deliberate transplantation.

**INTRODUCED BY NATURAL MEANS:** The plant is introduced, but the agent is considered not to be man. The actual agent may be anything other than man, including wind, current and animals for example, as vectors. The evidence leads to the conclusion that the plant has arrived without man's direct influence. If the agency is unknown, then it must be known not to be man, i.e. it is sufficient that there is evidence that the means of introduction was not man, without knowing the actual means of introduction.

**NO INFORMATION:** The plant is known to be introduced, but there is no information about the nature of its arrival, and the introduction could have been made by man or natural agency. In many cases the means of introduction will not be known, and this value should be applied.

**NOT APPLICABLE:** This data field does not in any way apply to the plant in this area because Data Field 3: Introduced Status, has the value NOT INTRODUCED or NO INFORMATION .

Data field 5: Cultivated Status

Whether or not the plant is cultivated in the area concerned.

| Value Suggested | Transfer Value |
| --- | --- |
| **CULTIVATED OUTDOORS** | C |

| | |
|---|---|
| **CULTIVATED INDOORS** | I |
| **ASSUMED** TO BE CULTIVATED | S |
| **DOUBTFULLY** CULTIVATED | D |
| **FORMERLY CULTIVATED (EXTINCT)** | E |
| **NOT** CULTIVATED | A |
| **RECORDED AS** CULTIVATED **IN ERROR** | F |
| **NO INFORMATION** | - |
| **NONE OF THE ABOVE** | U |
| **NOT APPLICABLE** | X |

*General*

This field is optional, and allows for detailed information to be recorded on the cultivation status of a taxon.

*Values Explained*

**CULTIVATED:** The plant is established in cultivation in the area concerned. The purpose of cultivation, be it for horticultural, economic, conservation, research needs etc, is not specified, nor is the type of cultivation described, such as whether the plant occurs in botanic gardens, is grown as a farm crop or grown in greenhouses and other indoor settings.

**ASSUMED TO BE CULTIVATED:** Assumed to be Cultivated in the area concerned. The assumption may be made, for example, on the basis of evidence about other species or neighbouring areas.

**DOUBTFULLY CULTIVATED:** There is doubt as to whether the status of the plant is cultivated or not in the area concerned. All records about the cultivated status of the plant in the area are in doubt.

**FORMERLY CULTiVATED (EXTINCT):** The plant was at one time cultivated, doubtfully cultivated or assumed to be cultivated in the area concerned and has become extinct in cultivation in this area, even though it may be extant elsewhere.

**NOT CULTIVATED:** The plant is definitely not cultivated (as defined above) in the area concerned.

**RECORDED AS CULTIVATED IN ERROR:** The plant has been recorded as Cultivated in the area concerned but all of those records have been disproved or discounted. A known fallacious record of cultivation must have been made, and it must be known that the plant is not cultivated in the area.

**NO INFORMATION:** There is no information as to whether the plant is Cultivated or not in the area concerned.

**NONE OF THE ABOVE:** This is included so as to identify any situations that may arise which cannot be covered by the above values.

**NOT APPLICABLE:** This data field does not in any way apply to the plant in this area because Data Field 1: Occurrence, has the value ABSENT.

## *DISTRIBUTION COMPLETENESS INFORMATION*

### Data Field 6: Area Distribution Completeness

Whether or not the complete distribution for the plant in the area in question is recorded.

| Value Suggested | Transfer Value |
|---|:---:|
| **DISTRIBUTION COMPLETE** | C |
| **DISTRIBUTION INCOMPLETE** | I |
| **NOT KNOWN WHETHER DISTRIBUTION COMPLETE** | U |
| **NOT APPLICABLE** | X |

*General*

This data field is designed to be used specifically with a hierarchical system for managing plant-area records in the database, such as the TDWG Standard No. 2 *World Geographical Scheme for Recording Plant Distributions* (Hollis & Brummitt, 1992) under which a 4-tiered hierarchy is defined. Under this scheme, level 4 comprises Basic Recording Unit's (BRUs). These BRUs are grouped to fit level 3 Botanical countries, which in turn are grouped into level 2 countries and regions. Level 1 defines the nine continents of the world.

The plant-area record to which this data field is applied should encompass other plant-area records in the database. An example, following the TDWG Standard 2 scheme, is the application of this data field to a level 3 record for India when the

database also contains records for individual level 4 areas such as Delhi and Goa. The main use of this data field might be to give an indication that, within the compass of the area concerned "All know areas" means areas in the sense used by the database, such as states within a country, not, for example, every site.

The level of detail to which the database using this data field appertains is important. If, for example, it is a country record, and the database also holds state records for that country, the value **DISTRIBUTION COMPLETE** can be applied when all states in which the plant occurs have been recorded for the country. At the lowest level of geographic detail in any database, often the state level, the value **DISTRIBUTION COMPLETE** will automatically apply to the record as soon as the information of its occurrence is recorded.

Regional databases could make use of an area record for a region, which if given a value for **Area Distribution Completeness** would therefore represent distribution completeness and therefore endemism for the compass of the database. More than one region may be covered by any database, and this approach allows for this situation. It also allows further refinement of the World Distribution Completeness category, but need not be implemented unless required.

*Values Explained*

**DISTRIBUTION COMPLETE:** Within the area covered by this plant-area record, all area records for the plant have been entered.

**DISTRIBUTION INCOMPLETE:** Within the area covered by this plant-area record, it is known that not all area records for the plant have been entered.

**NOT KNOWN WHETHER DISTRIBUTION COMPLETE:** Within the area covered by this plantarea record, it is not known if all the area records for the plant have been entered.

Data Field 7: World Distribution Completeness

Whether or not the plant-area records in the database represent the complete world native distribution for the plant.

| Value Suggested | Transfer Value |
|---|---|
| **DISTRIBUTION COMPLETE** | C |
| **DISTRIBUTION INCOMPLETE NOT KNOWN WHETHER** | I |

| DISTRIBUTION COMPLETE | U |
|---|---|

*General*

This field summarises the information in all the distribution records for a particular taxon,

and allows the area of endemism for each taxon to be calculated.

The use of this data field differs fundamentally from that of the other six because this data field links ONLY to the plant taxon and NOT to every plant-area record. It indicates the completeness of the area/distribution records for any one plant in a database and hence enables the plant to be identified as an endemic to one or more areas. It may be applied when all or part of the distribution of a species is covered in the database, and does not require the database to hold records of the entire world.

For example, if a plant occurs in only two known areas in the world, and they are both recorded in the database then **World Distribution Completeness** is set to the value **DISTRIBUTION COMPLETE**. This infers that the plant is endemic to those two areas. If, however, a plant occurs in many areas, only some of which are included in the database, then **World Distribution completeness** is set to **DISTRIBUTION INCOMPLETE**; thus the plant can be described as a non-endemic for the areas covered by the database.

This data field can only hold the value **DISTRIBUTION COMPLETE** if all values in Data Field 6: **Area Distribution Completeness** also have the value **DISTRIBUTION COMPLETE**.

However, even if all values in data field 6 - **Area Distribution Completeness are** set to have the value **DISTRIBUTION COMPLETE**, this does not necessarily mean that Data Field 7: **World Distribution Completeness** automatically has a value of **DISTRIBUTION COMPLETE**, unless the database concerned covers the entire world. For example, distribution records for a European database may be complete, but a taxon may occur in other areas outside Europe and hence not covered in this particular database. In this example, flagging **World Distribution Completeness** as **DISTRIBUTION INCOMPLETE** makes it evident that the taxon in question is not endemic to the area covered by the database.

The scale of the area records is important to this data field. If the database manages, for example, country records, then the value **DISTRIBUTION COMPLETE** can be applied when all countries in which the plant is known to occur have been recorded. It is not necessary in this case to have all the records for states or smaller areas in addition.

This data field applies to the native status of the plant, and may be assigned the value **DISTRIBUTION COMPLETE** when all known native records have been entered into the database. It is not relevant if the plant has a wider introduced or cultivated range.

*Values Explained*

**DISTRIBUTION COMPLETE:** Within the database, all area records for the plant have been
entered. The records cover the entire known range of the plant.

**DISTRIBUTION INCOMPLETE:** Within the database, it is known that not all area records for the plant have been entered.

**NOT KNOWN WHETHER DISTRIBUTION COMPLETE:** Within the database, it is not known if all the area records for the plant have been entered. Any number of area records (from zero upwards) may exist in the database.

The use of this field allows for the concept of "endemism" to be managed. For example:

*Arabis stricta* Hudson is known from Spain, France and Great Britain. It is therefore endemic to Europe, where Europe is the smallest defined area that contains its entire distribution.

1. The database holds a record for *A. stricta* Hudson in Spain.

The database holds a record for *A. stricta* Hudson in France.

The database holds a record for *A. stricta* Hudson in Great Britain.

2. The record for A. *stricta* Hudson has the value **DISTRIBUTION COMPLETE** in Data Field 7: World Distribution Completeness.

3. Therefore, since there can be no further records of *A. stricta* Hudson, it is endemic to the areas shown.

4. If the database has a hierarchical geographical system, and Europe is the area record which contains Spain, France and Great Britain, *A. stricta* Hudson is therefore endemic to Europe.

5. The values in Data Field 6: Area Distribution Completeness, for the records for Spain, France and Great Britain, are irrelevant to endemicity. These data fields need only be used if the database stores area records within any one or more of these countries, such as countries of Great Britain.