The University of Auckland and Landcare Research NZ Limited
Tamaki Campus
Private Bag 92019, Auckland, New Zealand
Tel: #64 9 3737 599 x 85210
Fax: #64 9 3737 042

Consultant's report to the Convention on Biological Diversity on development of a Draft Species Profile Schema for the Global Invasive Species Information Network (GISIN)

Consultants:
Michael Browne
IUCN SSC Invasive Species Specialist Group, Centre for Biodiversity and Biosecurity.
Jerry Cooper
Landcare Research NZ Limited, Manager GBIF New Zealand Node, Centre for Biodiversity and Biosecurity

INTRODUCTION

Exchanging data, information and knowledge about invasive species is a critical component of our response to the biological invasions. The Invasive Alien Species Profile Schema (IAS-PS) described in this report proposes preliminary draft standards for the exchange of invasive species information in the form of fact sheets, projects and case studies, expertise and checklists via the Global Invasive Species Information Network (GISIN). It supports the exchange of data elements important to invasive species science such as native and non-native status, pathways of spread, host and parasitic organisms observed, impact, invasiveness, control techniques used. The IAS-PS does not present schemas for data elements such as observations and survey data and bibliographies, but instead references existing schemas for the invasive species community to evaluate and adopt wherever possible.

The purpose of the IAS-PS is to facilitate data exchange and it should not serve as a template for information system design. It has been developed as an annotated XML-Schema document. We stress that the XML schema, and the suggested structure it encapsulates, will require further technical development. It is intended that the IAS-PS become the subject of discussion within the invasive species community and that an appropriate TDWG (Taxonomic Database Working Group) standard working group will take development of the proposed draft through a formal standards adoption process. Explanations of the main elements of the IAS-PS are provided in Appendix 1.

TERMS OF REFERENCE

The terms of reference for the consultants were to develop a data exchange standard for Alien Invasive Species Profiles by consulting main actors, comparing and evaluating different schemas and data formats and to propose a core set of data elements. This has been achieved and recommendations have been made for further development.

Rather than compare and contrast terminology used by the invasive species community, as outlined in the terms of reference, we identified high-level elements that can be driven by standard terminologies/dictionaries (e.g. ISO standards) and proposed some universal vocabularies for discussion. We concluded that resolving the issues surrounding use of the term 'invasive' could not be achieved within the scope of the current IAS-PS. The IAS community will need to discuss further whether a universal definition of the word 'invasive' for GISIN is desirable, or whether it is sufficient to attach the working definitions used to the data being exchanged, as provided for in the schema. In addition, we have ensured that the IAS-PS contains data elements for impacts, abundance, distribution and population dynamics, which can underpin different uses of the word 'invasive'

APPROACH

In March 2005 we circulated a discussion paper that articulated a vision for the GISIN, identified possible roles that an IAS-PS might play in GISIN and raised some of the complex issues that would need to be addressed in the IAS-PS.

Responses to the discussion paper included comments on the importance of biostatus, life history and links to reputable taxonomic authorities, the challenge in constructing and agreeing upon controlled vocabularies and creating crosswalks, the importance of awareness raising and providing onward links to other sources of information and the use of 'virtual fact sheets' to identify where information is weak or lacking. Other issues were outside the scope of the IAS-PS and they included the importance of multiple languages, sharing material on IAS policy, and sharing information on the susceptibility of specific ecosystems.

The Consultants then began developing the draft IAS-PS by exploring a larger number of online IAS information sources and classifying data elements they use. This was the best method to ensure that the IAS_PS would cope with the enormous variety and complexity of IAS information, given that invasive species range from micro-organisms to large plants and animals and that they occur in all realms. We used local, national, regional and global sources of information that describe marine, freshwater and terrestrial invasive species, and we included species that impact biodiversity and agricultural systems. See Appendix 2 for a list of online databases and other resources that we examined.

CHALLENGES

Managing expectations

In developing an IAS Profile schema we have not addressed how the schema should be used by the GISIN community. The vision and strategy for that network is not yet clear. GISIN may simply provide searchable access to equivalent data types from a distributed network of data providers. But how will end users then filter multiple and overlapping data for the items of interest? Alternatively, the GISIN network of data providers may feed data into in a centrally coordinated clearing house of IAS data where duplication and overlapping information is managed. That management mechanism is beyond the scope of this data exchange schema.

In spite of the reference to species profiles in the IAS-PS's title, there may be expectations that the IAS-PS will support all types of IAS information. In fact, the IAS-PS in its current form only supports the following data types:

- Fact sheets/profiles
- Non-native and invasive checklists
- Positional, relational, bioregional, biogeographic data
- Distributional data
- Management projects and case studies
- Expertise

Existing schemas

Where available, existing schemas should be used for data types such as; agents (experts), observation and survey data, specimen data, bibliographies, identification or diagnostic information, spatial data and images. For example, the Darwin Core is minimally adequate for one-off, opportunistic observation/collection data (as most museum/collection/society oriented records are), but isn't appropriate for structured monitoring data where methodology is critical and variable. The IAS_PS has a 'place holder' for primary data indicating that it should be shared via the Darwin Core, ABCD or Structured Survey Data schemas.

Where existing schemas are not available or not suitable for invasive species information, further work is required. Each invasive species specialist community may need to develop their own specific extensions for use within the IAS Profile Schema. For example, ant data requires descriptions for % bare ground, % leaf litter cover, % coarse woody debris cover, canopy structure, canopy height and % canopy cover.

Primary and secondary data

IAS data is fundamentally different in character to data shared within existing networks like GIBIF. GBIF has a focus on primary data and, in general terms, there are existing mechanisms for defining the boundaries of the data being shared, for example through a unique collection accession number. This delineation and identification of data facilitates the identification of duplicates. IAS data is generally a mix of primary data (for example collection/observation distribution information), collated together with descriptive, management and other kinds of data, which are either being reported directly or reported from the literature. IAS Profiles are therefore generally a synthesis of primary and derived data from many potential sources.

Documenting sources

Given that much of the content that will be shared in GISIN is summarised information from secondary sources, is it vital that the sources of information be fully documented. The task of specifying and tracking the provenance of data and

relationships between data components, even within a single profile, is difficult to capture within a data exchange schema, and often not fully recorded within the source IAS databases. In the draft schema presented here we have adopted a pragmatic approach to this attribution problem.

Key information elements are accompanied by elements that specifically record attribution. At a higher level we also suggest an indirect method where attribution may be enumerated independently of each and every element but linked to them by keywords. For example impact data is carried by an element deep within the schema, but the provenance of those data, e.g. a publication, may be encoded separately and linked to it via an 'impact' keyword.

A note on XML-Schema

We have chosen to model the data exchange profile using XML-Schema. This is currently the most common method for defining data exchange standards. XML-Schema allows packets of information to be encoded as a hierarchy of data elements. Some existing schema representations of data exchange standards, for example Darwin Core, are very simple and use only a flat enumeration of data elements with no hierarchical depth or repeated elements. Other schemas use a richer information model, for example ABCD. The nature of IAS Profiles necessitates the use of a richer hierarchical model with repeated data elements. At first sight the draft schema we have developed may appear overly complex. However, IAS databases not only vary considerably in the range and provenance of the data they contain, but they also vary considerably in the degree of atomisation of these data. For example, in one database the distribution data for a species may be a simple block of text, whilst in others it may be a richer relational structure capturing information about ISO standard country names, municipalities etc. The schema presented here is intended to be applicable to as many data providers as possible. Data should be encoded by the schema at the appropriate level of atomisation. A relatively simple IAS database structure would therefore only be expected to populate relatively higher elements in the schema, whereas those databases with richer data models should use deeper levels of atomised elements.

Species names:

The name of an organism does not always uniquely define the taxon to which it nominally refers. It is common for a name to be used by different authorities to refer to different, overlapping and distinct biological entities. Within the IAS-PS we include the option to reference the future GBIF Electronic Catalogue of Names (ECAT) Globally Unique Identifiers (GUID) for taxonomic concepts. If IAS-PS data providers maintain significant and authoritative taxonomic data, then they should consider becoming GBIF ECAT data providers, or at least linking their existing nominal taxonomic data to GUIDs within the developing authoritative catalogues such as ECAT, and the Catalogue of Life.

Users of the IAS-PS can record a scientific name and author and the reference they are using for taxonomy. They can select a term from the life form lookup and record kingdom, family name and common names. The data element linking a defined/nominal taxon concept to the IAS-PS is the appropriate Taxon Concept Schema (TCS) GUID. This external reference will be used in the future to resolve the various related name/taxon concepts (e.g. synonyms). However, it should be noted that, at the time of writing, GBIF's ECAT programme does not yet provide a TCS-

mediated data provider network for taxonomic data, or the necessary resolution service for TCS GUIDs. When this support becomes available, users will be able to choose scientific names from a pick list.

The GISIN infrastructure

GISIN is intended to be a platform for sharing IAS information at a global level. The IAS-PS is a technical definition for implementing data sharing, but the infrastructure in which it would operate has not yet been defined. To provide a useful service, GISIN requires a coordinating body, operational infrastructure, a registry of participating databases, an indexing and clearing house service. If at least one portal provides an integrated view of the data derived from a GISIN network it will be able to demonstrate the power of the network e.g. by maintaining a definitive checklist of all known invasive species for those conducting risk assessments. There are opportunities at this level for collaboration between GBIF, GISIN and other relevant agencies.

To avoid duplication and strengthen infrastructure, GISN should take advantage of TDWG's formal standards adoption process for developing this draft standard further. In addition any GISIN network should be aligned with the planned development of the communication software for biological data-sharing, for example DiGIR (Distributed Generic Information Retrieval) and BioCASE (Biological Collection Access Service for Europe). If these software wrappers are developed generically enough, the IAS-PS can be one of a number of data exchange standards supported within these data sharing communication protocols.

Geospatial data

The current IAS-PS provides for the definition of Web Mapping Service. Using GISIN to integrate spatial and non-spatial data within GIS mapping portals is one potential role of a central 'clearing house', but all that is outside the scope of the current schema.

Terminology

The IAS-PS contains proposed standard vocabularies for Life Form, Realm, Habitat, Country name (ISO3166), Biostatus (occurrence, status, invasiveness), Impact (Type, Mechanism, Outcome) Introduction and Dispersal (Type, Date, Cause, Route, Vector) (see Appendix 4 for proposed standard vocabularies). Use of these standard vocabularies will allow high-level information in these categories to be readily exchanged. We recommend that the invasive species community debate the usefulness of the standard vocabularies proposed in the IAS-PS. Additional vocabularies should be proposed and discussed, bearing in mind that they are more likely to be used if they remain relatively simple.

Foremost among terminology issues, is use of the word 'invasive'. In its widest sense, 'invasive' can refer to any organism that is a nuisance, or it can be restricted to organisms that negatively impact biodiversity. Invasiveness is often expressed in terms of 'impacts' for animals, and in terms of distribution and abundance for plants. Use of the word 'invasive' has developed independently in different sectors like agriculture and environment, and in the key international instruments (CBD, IPPC,

IMO, etc.)[1] that address invasive alien species, reflecting their different mandates. Ecological concepts and legal concepts are not always easy to reconcile, and research environments and management / regulatory environments have different requirements.

We recommend that the IAS community discuss further whether a universal definition of the word 'invasive' for GISIN is desirable, or whether it is sufficient to attach the working definitions used to the data being exchanged, as provided for in the schema. In addition, we have ensured that the IAS-PS contains data elements for impacts, abundance, distribution and population dynamics, which can underpin different uses of the word 'invasive'

DATA TYPES SUPPORTED BY THE IAS-PS

Fact Sheets/Profiles

Fact sheet type information includes descriptions, biology, climate and habitat, native and introduced range, pathways, impacts, management and much more. Fact sheets typically contain a mixture of generic and location-specific information. Some data, such as date of arrival or abundance, is clearly location-specific, but often the link between the information and a location is not stated explicitly, and the spatial context of statements is often unclear. In the IAS-PS, one response is to add geographical, thematic and taxonomic scope elements to the ABCD Metadata schema. The other is to provide a GeneralFacts / LocationSpecificFacts switch within the schema, and to record a location name. This is easily achieved at the fact sheet level in the current version of IAS-PS, but assigning separate 'facts' to different locations is a slower process. The community should discuss the costs and benefits of providing the GeneralFacts / LocationSpecificFacts switch option at all levels of the schema (profile level, element level, sub-element-level, sub-sub-element-level, etc.), since information at any level can potentially be location-specific.

Non-native and invasive checklists

Checklists are essentially lists of species recorded for a location (e.g. a country) for a purpose. Their purpose is usually to identify invasive species or non-indigenous species (that are causing some concern). Checklists are supported in the ReportedData component of the IAS-PS by using Taxon and Location elements, along with components of Biostatus from the LocationSpecificFacts element.

Positional, relational, bioregional, biogeographic data

The IAS-PS contains a number of extensions to the Darwin Core V2.0 for positional, relational, bioregional, biogeographic data. They include a LocationGUID to link to an external reference, AreaSize, Island, IslandGroup, MarineBioregion and FreshWaterBody (the last two elements include a Defined Schema sequence option).

Distributional data

Distributional data is supported in the ReportedData component of the IAS-PS by using Taxon and Location elements, and elements from Facts and LocationSpecificFacts.

Management projects and case studies

[1] Convention on Biological Diversity, International Plant Protection Convention, International Maritime Organisation

Information about management projects and case studies can be shared via the ProjectOrCaseStudy element, which can also be used for broader topics (e.g. case studies of impacts, introduction and dispersal). Taxon and Location elements are required and elements from Facts and LocationSpecificFacts can be used for background information. The many sub-elements of ProjectOrCaseStudy can carry information about aims and achievements, dates, activities, costs and conclusions.

Expertise
The exchange of expert's details, including their contact details, affiliation and geographic and taxonomic area of expertise is supported in the current version of the IAS-PS.

TOPICS FOR DISCUSSION

We recommend that the community should discuss the need for a coordinating body, operational infrastructure, a registry of participating databases, an indexing and clearing house service. If at least one portal provides an integrated view of the data derived from a GISIN network it will be able to demonstrate the power of the network e.g. by maintaining a definitive checklist of all known invasive species for those conducting risk assessments. Using GISIN to integrate spatial and non-spatial data within GIS mapping portals is one potential role of a central 'clearing house', but all that is outside the scope of the current schema.

We recommend that the community should discuss opportunities for collaboration between GBIF, GISIN and other relevant agencies.

We recommend that the invasive species community debate the usefulness of the standard vocabularies proposed in the IAS-PS (Appendix 4). Additional vocabularies should be proposed and discussed, bearing in mind that they are more likely to be used if they remain relatively simple. In particular, the community should work towards agreement on a working definition of the word invasive for GISIN.

We recommend that the community should discuss the costs and benefits of providing the GeneralFacts / LocationSpecificFacts switch option at all levels of the schema (profile level, element level, sub-element-level, sub-sub-element-level, etc.), since information at any level can potentially be location-specific.

ACKNOWLEDGEMENTS

Thanks to the CBD for the opportunity to develop an IAS-PS and for posting the March 2005 Discussion Paper on their website. Thanks to all the contributors who provided valuable suggestions. Thanks to NBII for their support and the list of online IAS information sources, and to GBIF for its Species Banks inventory. Thanks in particular for the gems that we found as we explored the world's online IAS information sources and to all the people who contributed their knowledge and support.

APPENDICES

Appendix 1. Explanation of IAS-PS elements

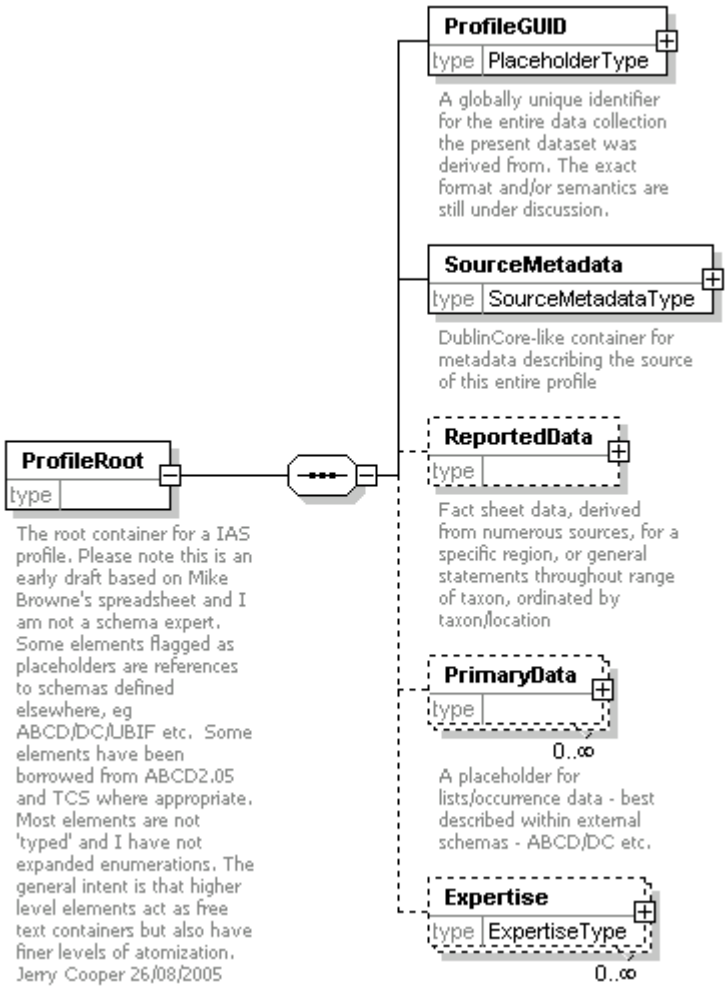
Introduction

Most elements in the IAS-PS have annotated comments that explain their purpose. This section of the report explains the way elements link to each other. Since the IAS-PS was originally intended to facilitate the exchange of fact sheets, that is the main focus of this section. Data types not covered in depth here include spatial data, images, diagnostic keys, bibliographies, specimen data, and observation, survey and collection data.

Elements in the IAS-PS are intended to carry free text so that the schema can support data at any level of atomisation. However, users should be encouraged to atomise their data as much as possible in order to make it more readily accessible to others.

Please note that this schema is an early draft. It does not have extensive data typing applied or explicit enumerations for elements using controlled vocabularies etc.

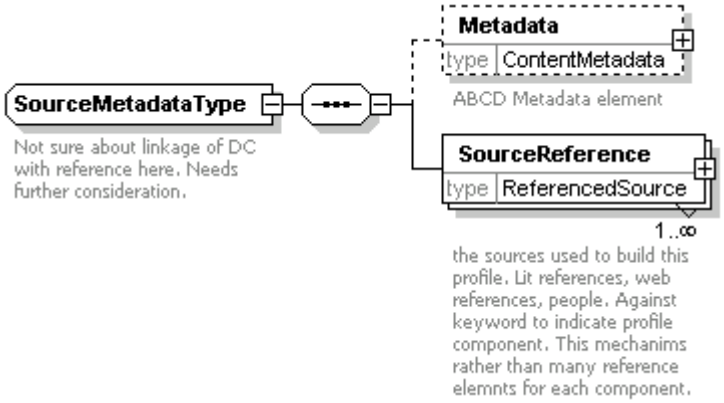
Fig.1 The Profile root



The ProfileRoot (*Fig. 1*) contains a GUID element, which describes the institution that generates or curates the data, a SourceMetadata element, which describes the

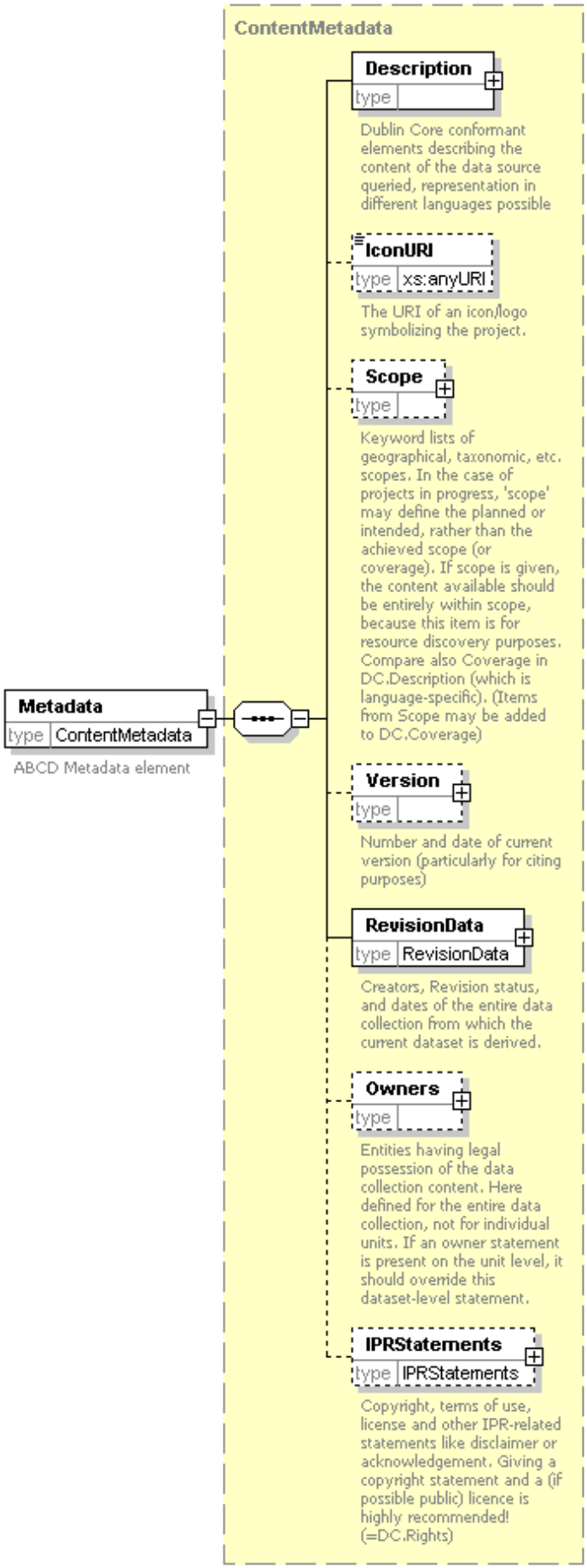
source of the actual data set (e.g. a fact sheet) being shared, a ReportedData element where facts sheet, checklist, project and case study data can be shared. In addition there is a PrimaryData element, which is simply a placeholder for observation, survey and collection data, and an Expertise element for sharing details about experts.

Fig. 2 The SourceMetadataType element



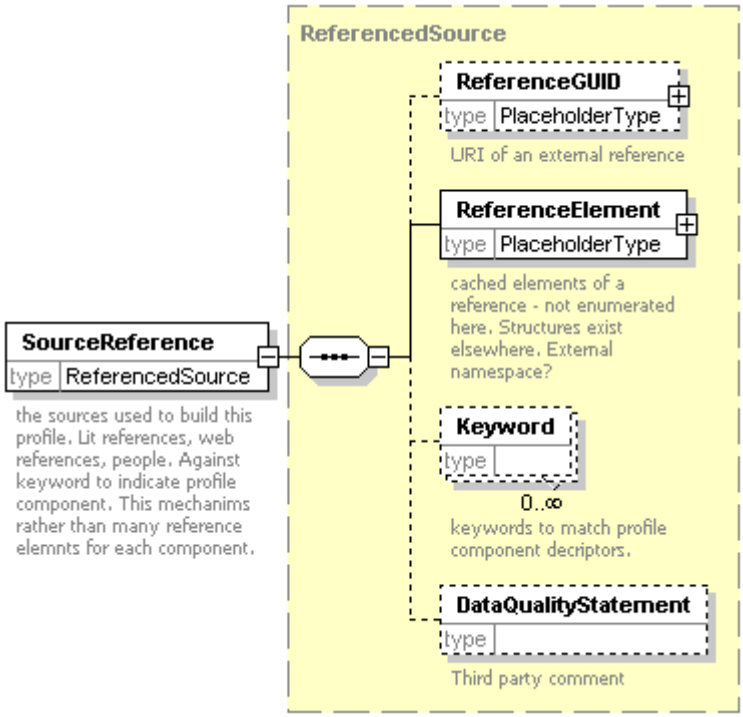
The SourceMetadataType element (*Fig. 2*) contains a Metadata element (*Fig. 3*), which carries contextual information about the data being exchanged and a SourceReference element (*Fig. 4*), for the references used in the dataset (e.g. fact sheet).

Fig. 3 The Metadata element



The Metadata element (*Fig. 3*) uses the ABCD content metadata standard for content description and language, logo, scope (geographic and taxonomic), version, revision status, owner and IPR statements.

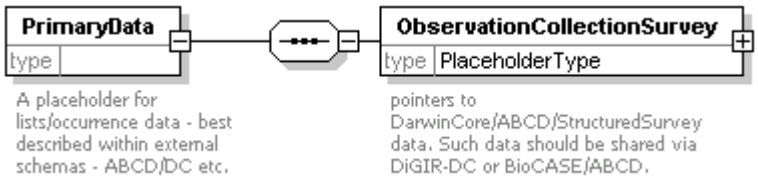
Fig. 4 The SourceReference element



The SourceReference element (*Fig. 4*) allows users to document sources and to record data quality statements. It can carry a link to an external reference (Reference GUID), cached elements of other reference schemas (ReferenceElement), keywords to link the reference to particular schema elements, and data quality statements.

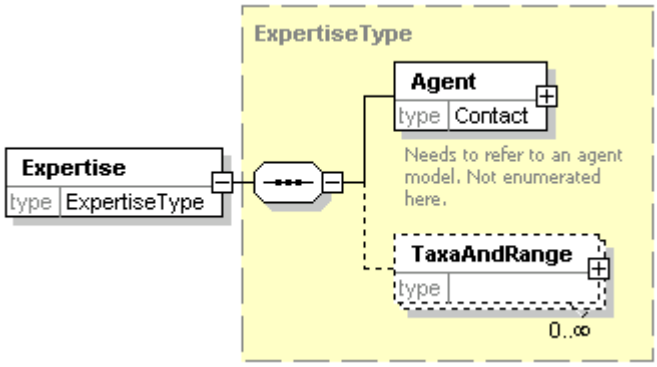
The ReportedData element, which carries fact sheet type information, is described below (*Fig 7*).

Fig. 5 The PrimaryData element



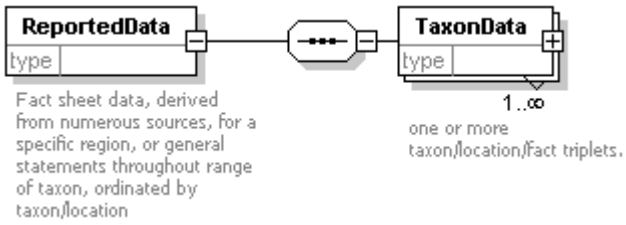
The PrimaryData (*Fig.5*) element is simply a placeholder for observation, survey and collection data which is best enumerated within a different schema (Darwin core, ABCD, structured survey data etc.).

Fig. 6 The Expertise element



The Expertise element (*Fig. 6*) contains Agent and TaxaAndRange sub-elements. The Agent element contains Organisation, Person, Roles, Addresses, TelephoneNumbers, EmailAddresses, URIs and LogoURI sub-elements. The TaxaAndRange element contains Range and Taxa sub-elements.

Fig. 7 The ReportedData element



The ReportedData element (*Fig. 7*) links combinations of Taxon, Location and Facts elements via the TaxonData element (*Fig. 8*). See Fig. 9 for the sub-elements of the Taxon element.

Fig. 8 The TaxonData element

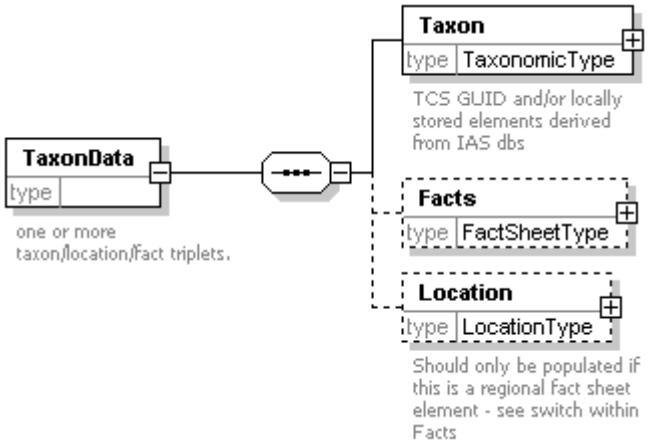
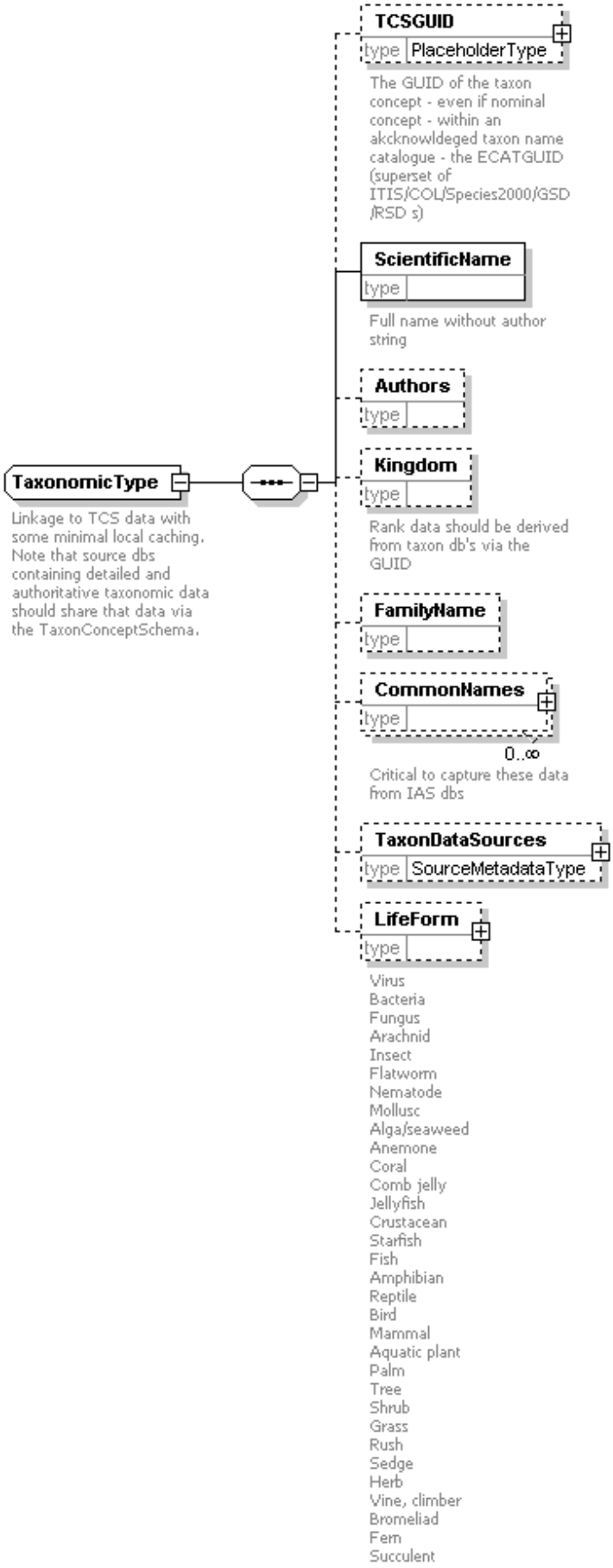
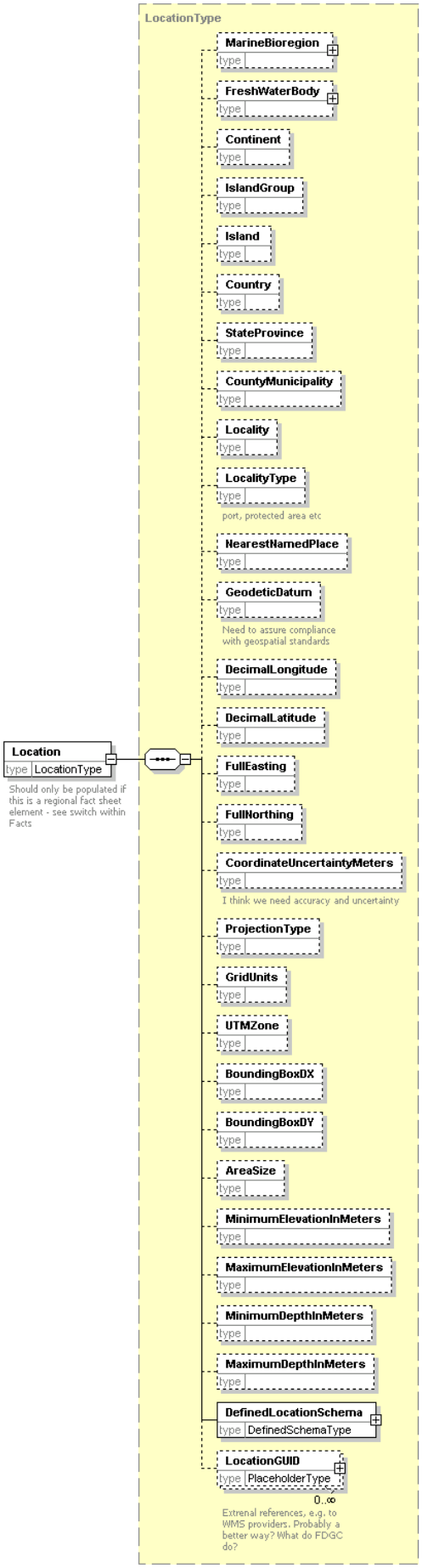


Fig. 9 The TaxonomicType element



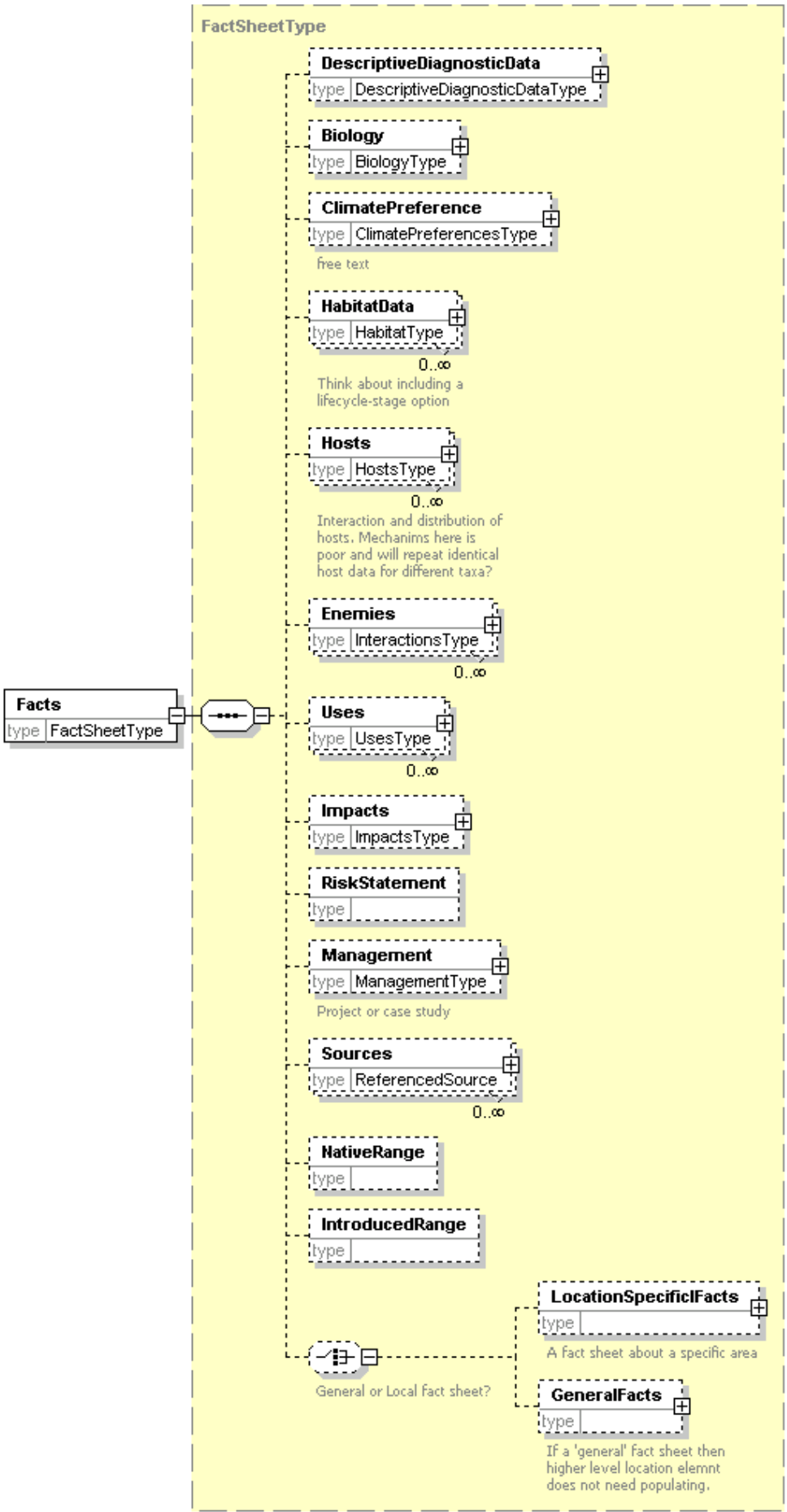
The TaxonomicType element (*Fig. 9*) contains standard Darwin Core elements such as TCSGUID, ScientificName, Authors, Kingdom, FamilyName, CommonNames, TaxonDataSources, and LifeForm. The LifeForm element contains a lookup that will support queries using simple terms like fish, insect, grass or reptile, and a DefinedSchemaType sequence to support information such as “Ardisia elliptica is a geophyte according to the Raunkiaer system”. The TCSGUID ensures continuity when an organism’s name is changed. The CommonNames element allows the language and location of use of a common name to be recorded against each common name.

Fig. 10 The Location Type element


LocationType
MarineBioregion
type
FreshWaterBody
type
Continent
type
IslandGroup
type
Island
type
Country
type
StateProvince
type
CountyMunicipality
type
Locality
type
LocalityType
type
port, protected area etc
NearestNamedPlace
type
GeodeticDatum
type
Need to assure compliance with geospatial standards
DecimalLongitude
type
DecimalLatitude
type
Location
type LocationType
Should only be populated if this is a regional fact sheet element - see switch within Facts
FullEasting
type
FullNorthing
type
CoordinateUncertaintyMeters
type
I think we need accuracy and uncertainty
ProjectionType
type
GridUnits
type
UTMZone
type
BoundingBoxDX
type
BoundingBoxDY
type
AreaSize
type
MinimumElevationInMeters
type
MaximumElevationInMeters
type
MinimumDepthInMeters
type
MaximumDepthInMeters
type
DefinedLocationSchema
type DefinedSchemaType
LocationGUID
type PlaceholderType
0..∞
Extrenal references, e.g. to WMS providers. Probably a better way? What do FDGC do?

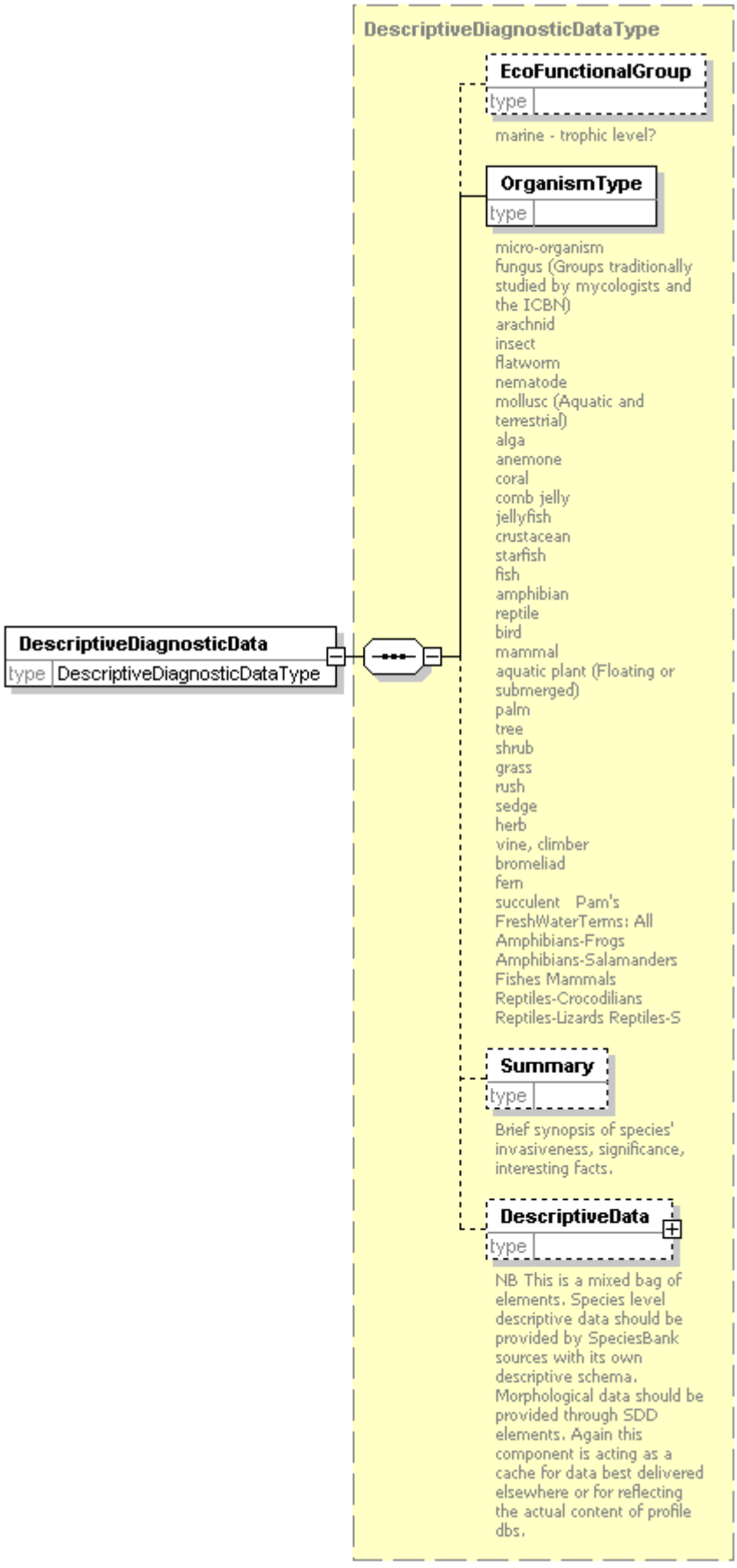
The Location element (*Fig. 10*) contains standard Darwin core elements and extensions. It includes MarineBioregion, FreshWaterBody, Continent, IslandGroup, Island, Country, StateProvince, CountyMunicipality, Locality, LocalityType, NearestNamedPlace, GeodeticDatum, DecimalLongitude, DecimalLatitude, FullEasting, FullNorthing, CoordinateUncertaintyMeters, ProjectionType, GridUnits, UTMZone, BoundingBoxDX, BoundingBoxDY, AreaSize, MinimumElevationInMeters, MaximumElevationInMeters, MinimumDepthInMeters, MaximumDepthInMeters, DefinedLocationSchema and LocationGUID. The DefinedLocationSchema element contains DefinedSchemaName, DefinedSchemaValue, and DefinedSchemaReference sub-elements.

Fig. 11 The FactSheetType element



The FactSheetType element (*Fig. 11*) contains elements and sub-elements one typically finds in fact sheets. Facts include DescriptiveDiagnosticData, Biology, ClimatePreference, HabitatData, Hosts, Enemies, Uses, Impacts, RiskStatement, Management, NativeRange, IntroducedRange and Sources. Explanations of each of these elements follow.

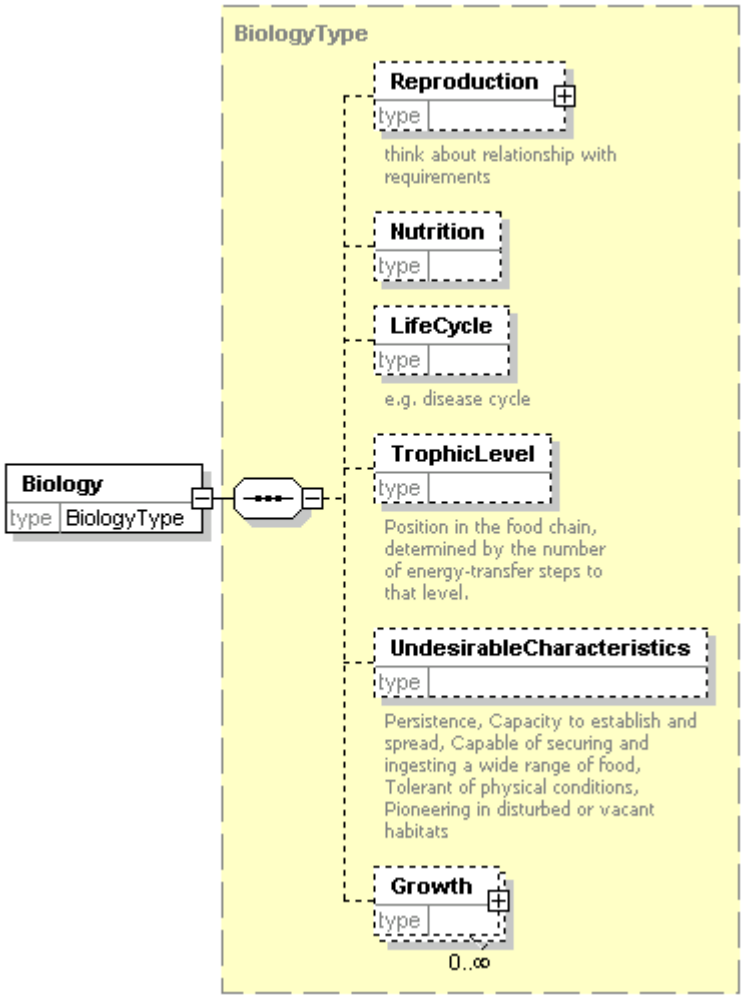
Fig. 12 The DescriptiveDiagnosticData element



The DescriptiveDiagnosticData element (*Fig. 12*) contains EcoFunctionalGroup, OrganismType, Summary, and DescriptiveData sub-elements. The DescriptiveData

element contains GlossaryURI, SpeciesBankURI, SDD-URI, ImageURI, MaximumSize, Phenology, KeyFeatures, AdultDescription, JuvenileDescription, References and Lookalikes sub-elements. The Lookalike element contains Remarks, LookalikeTaxon and ImageURI sub-elements. The LookalikeTaxon element has the same components as the 'global' TaxonomicType sub-element in *Fig. 9*.

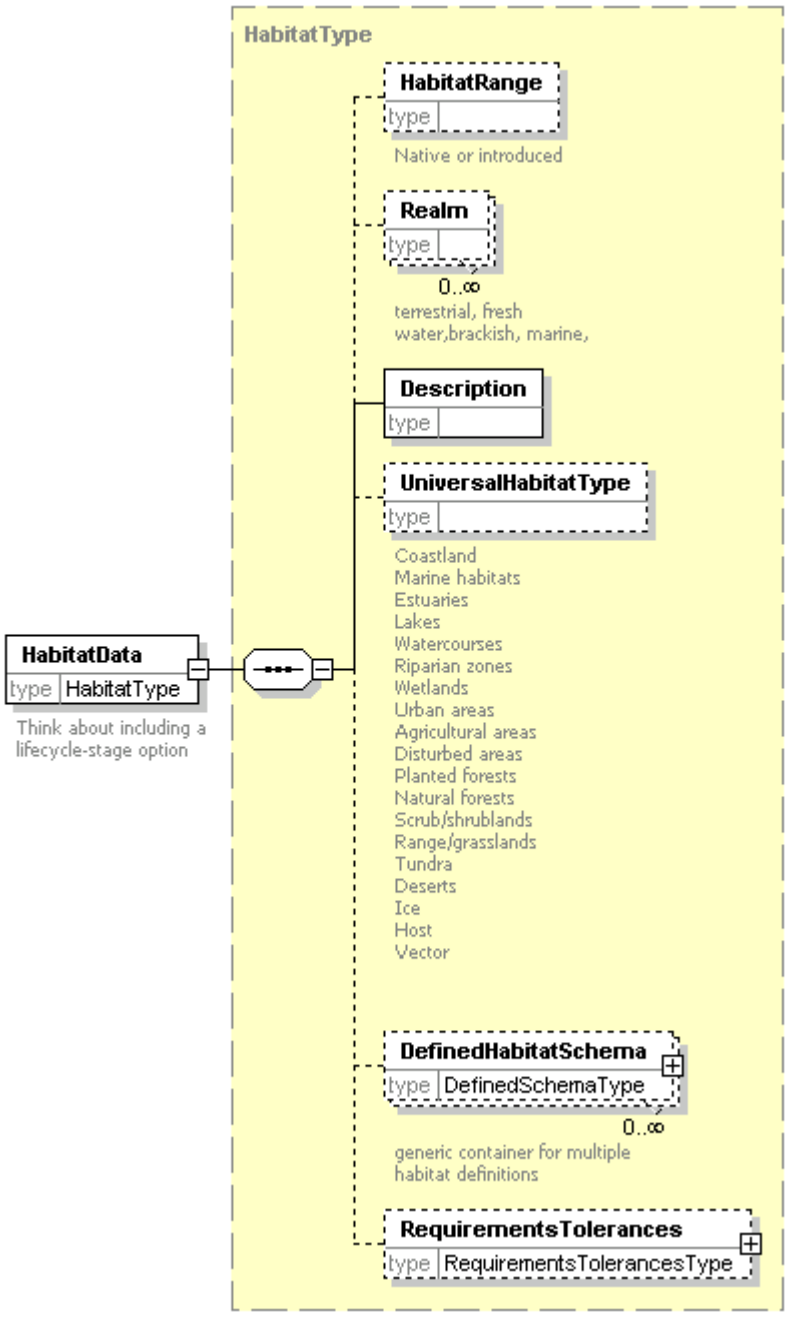
Fig. 13 The Biology element



The Biology element (*Fig. 13*), contains sub-elements for Reproduction (with 5 sub-elements; Temperature, Cues, Salinity, StrategyOrSystem, FecundityOrOutput), Nutrition, LifeCycle, TrophicLevel and UndesirableCharacteristics. The UndesirableCharacteristics element might include statements about persistence or time known in that location, capacity to establish and spread or to secure and ingest a wide range of food, tolerance of physical conditions, pioneering capacity in disturbed or vacant habitats, etc.

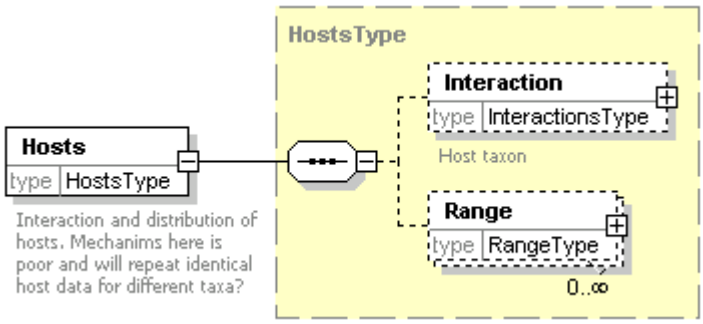
The ClimatePreference element contains a simple climate lookup (tropical, subtropical, temperate, boreal, polar, all).

Fig. 14 The HabitatData element



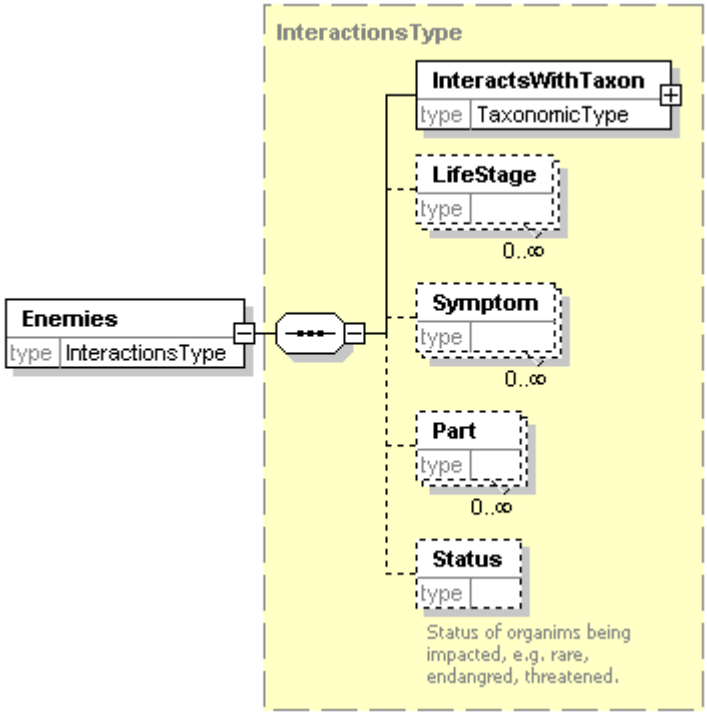
The HabitatData element (*Fig. 14*) contains Realm, Description, UniversalHabitatType, DefinedHabitatSchema, and RequirementsTolerances sub-elements. The Realm element has a lookup (terrestrial, freshwater, brackish, marine). The UniversalHabitatType element also has a lookup to facilitate searches using terms such as estuaries, wetlands and planted forests. The DefinedHabitatSchema element allows users to reference and use their own habitat classification systems. The RequirementsTolerances element contains Temperature, Precipitation, Elevation, Depth, Pedology, Salinity, Frost, Flooding, Drought, SolarRadiation, pH, DissolvedOxygen, WaterVelocity, Turbidity and Shading sub-elements. Some of these elements have sub-elements. This is an area for further development by those working on different invasive taxa.

Fig. 15 The Hosts element



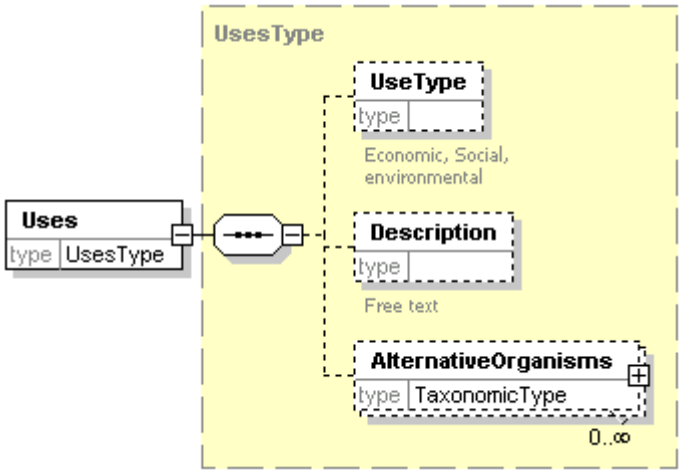
The Hosts element (*Fig. 15*) contains Interaction and Range (of the host) sub-elements. The Interaction element uses the global TaxonomicType element in *Fig. 9* to identify the host, and includes Lifestage, Symptom, Part and (threatened) Status (of the organism being impacted).

Fig. 16 The Enemies element



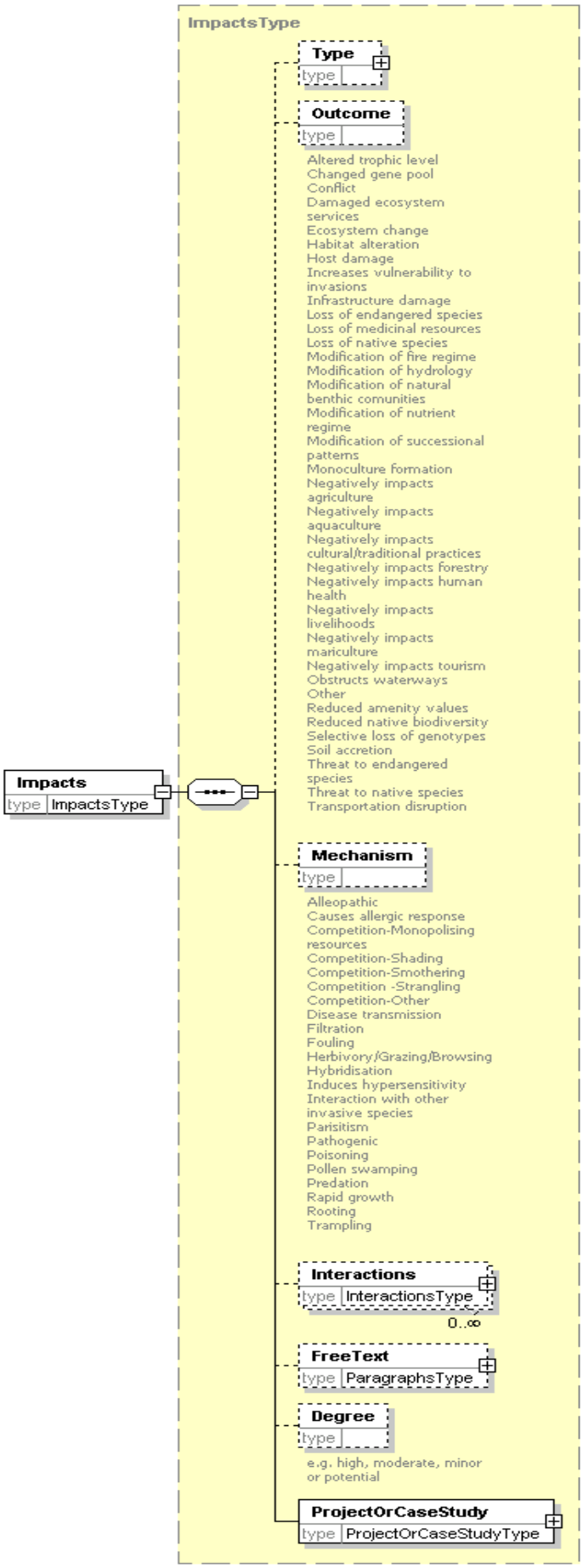
The Enemies element (*Fig. 16*) contains InteractsWithTaxon, LifeStage, Symptom, Part and Status sub-elements. The InteractsWithTaxon sub-element has the same components as the global TaxonomicType element in *Fig. 9.*

Fig. 17 The Uses element



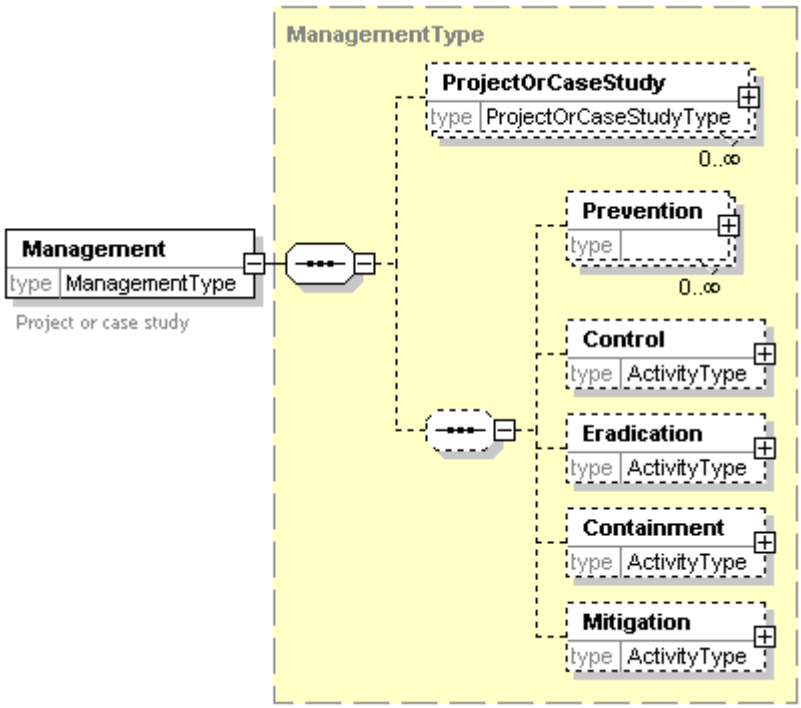
The Uses element (*Fig. 17*) contains UseType, Description and AlternativeOrganisms sub-elements and the AlternativeOrganisms sub-element has the same components as the ‘global’ TaxonomicType element in *Fig. 9*.

Fig. 18 The Impacts element



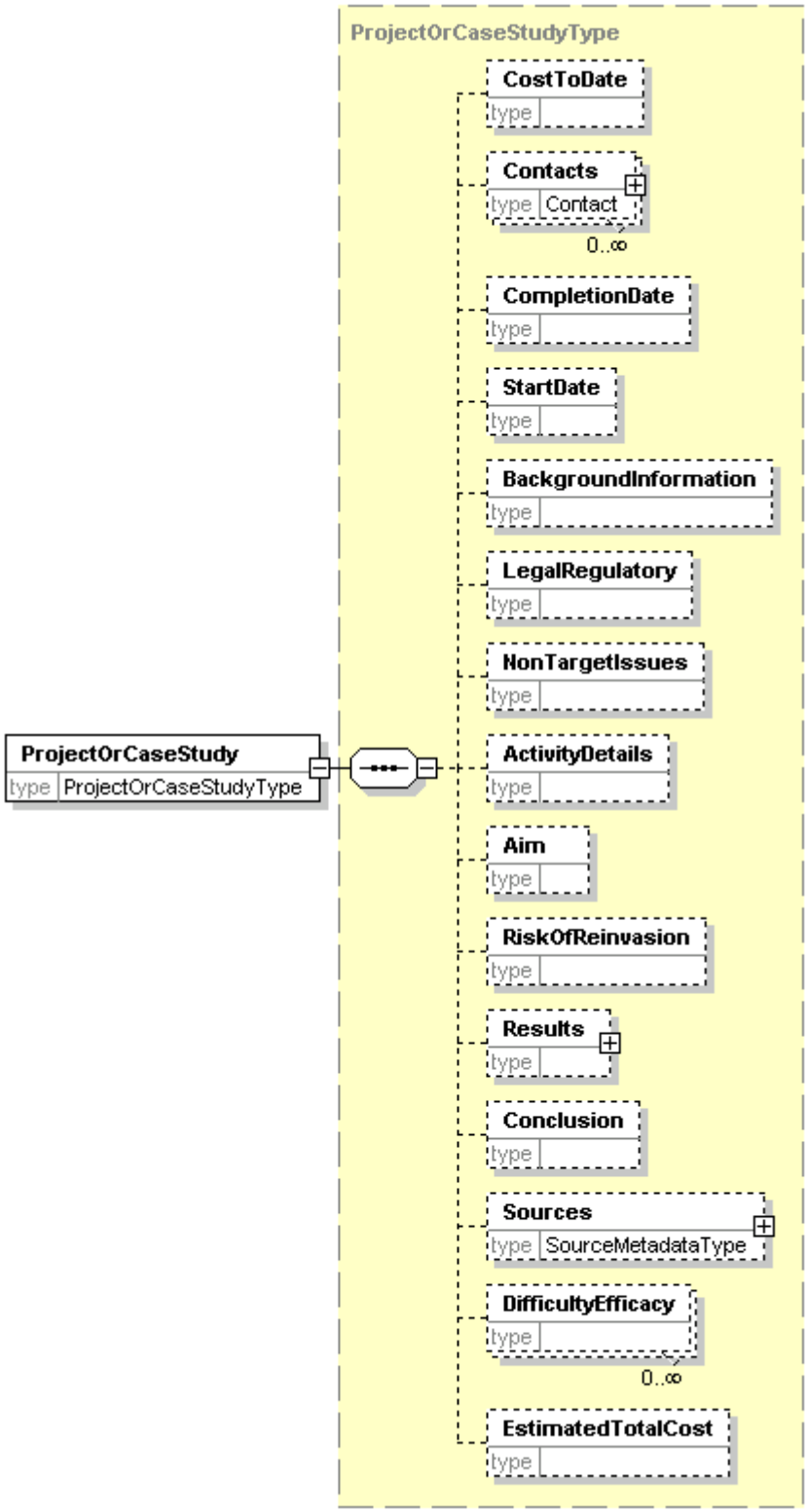
The Impacts element (*Fig. 18*) contains ImpactType, ImpactOutcome, ImpactMechanism, Interactions, Degree and ProjectOrCaseStudy sub-elements. The ImpactType element contains Environmental, HumanHealth, EconomicLivelihood and CulturalAmenity sub-elements, and the Environmental sub-element contains EcosystemProcesses, CommunityStructure, CommunityComposition and TaxonImpacted sub-elements. The TaxonImpacted element contains a Taxon sub-element, which has the same components as the 'global' TaxonomicType element in *Fig. 9*, and a ConservationSignificance sub-element. The Interactions element contains InteractsWithTaxon, LifeStage, Symptom, Part and Status sub-elements, and the InteractsWithTaxon element has the same components as the 'global' TaxonomicType element in *Fig. 9*. The ProjectOrCaseStudy element is described below.

Fig. 19 The Management element



The Management element (*Fig. 19*) contains a ProjectOrCaseStudy element for detailed records and Prevention, Control, Eradication, Containment and Mitigation sub-elements. The Preventative element contains IncursionDetection, RapidResponse, AwarenessRaising, and PostPreBorder sub-elements. The Eradication, Control, Containment and Mitigation elements contain identical Physical, Biological, Chemical, Cultural and IPM, sub-elements. The ProjectOrCaseStudy element is explained below.

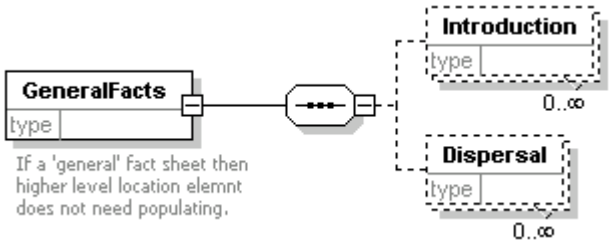
Fig. 20 The ProjectOrCaseStudy element



The ProjectOrCaseStudy element (*Fig. 20*) allows detailed information about projects and case studies to be shared. It is based on records collected for management projects but can also be used for case studies of impacts, introduction and dispersal.

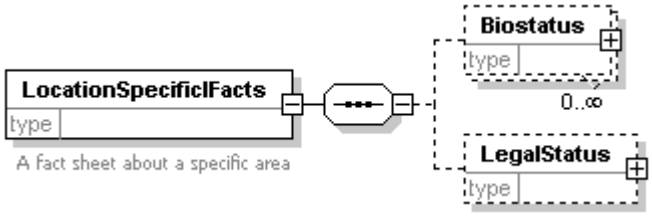
The Sources element uses the global ReferencedSource sequence (see *Fig. 4*) for documenting sources.

Fig. 21 The GeneralFacts element



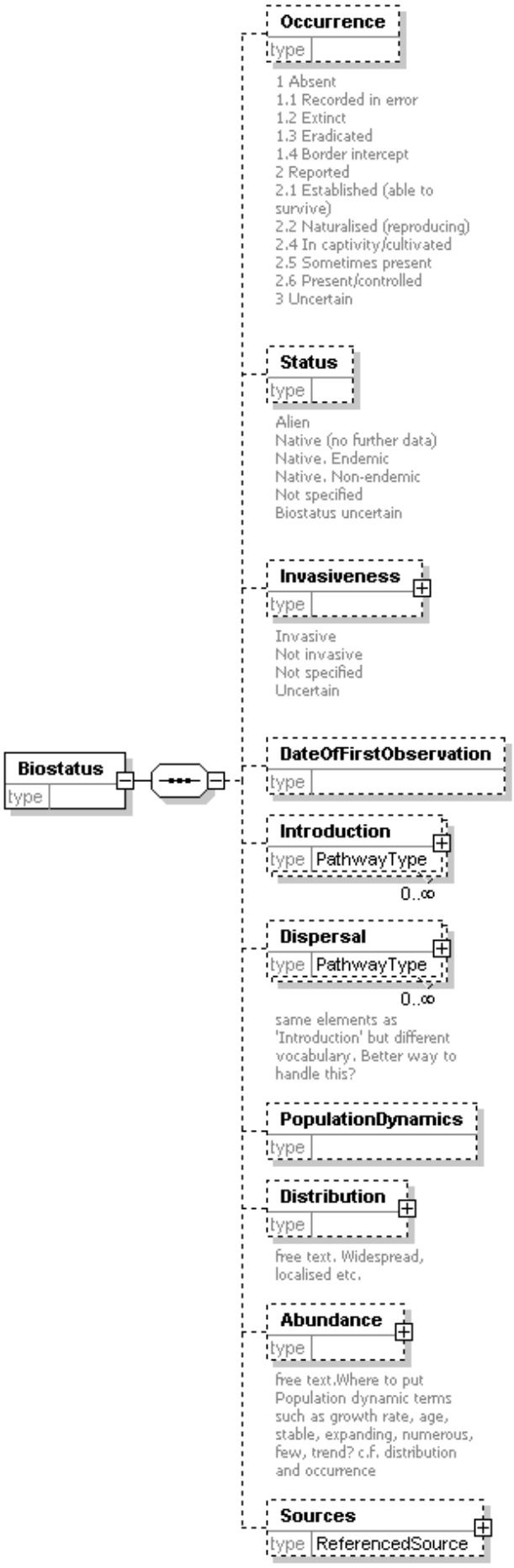
The final element amongst the Facts is a switch that allows the Facts to be identified either as either GeneralFacts (*Fig. 21*) or LocationSpecificFacts (*Fig. 22*). By choosing GeneralFacts, the user can add general information about Introduction and Dispersal, but if this information is location-specific, he/she must choose LocationSpecificFacts. In the current version of the IAS-PS, the only way to record some of the facts as GeneralFacts and others as LocationSpecificFacts is to make two separate 'Fact' entries. The community could usefully explore the costs and benefits of making this option more widely available at a lower level to cope with "mixed-content fact sheets".

Fig. 22 The LocationSpecificFacts element



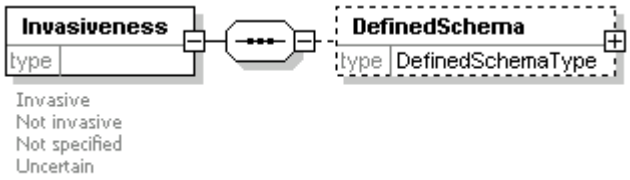
Where Facts are location-specific (*Fig. 22*), the user can record LegalStatus (using a user-defined schema) and Biostatus.

Fig. 23 The Biostatus element



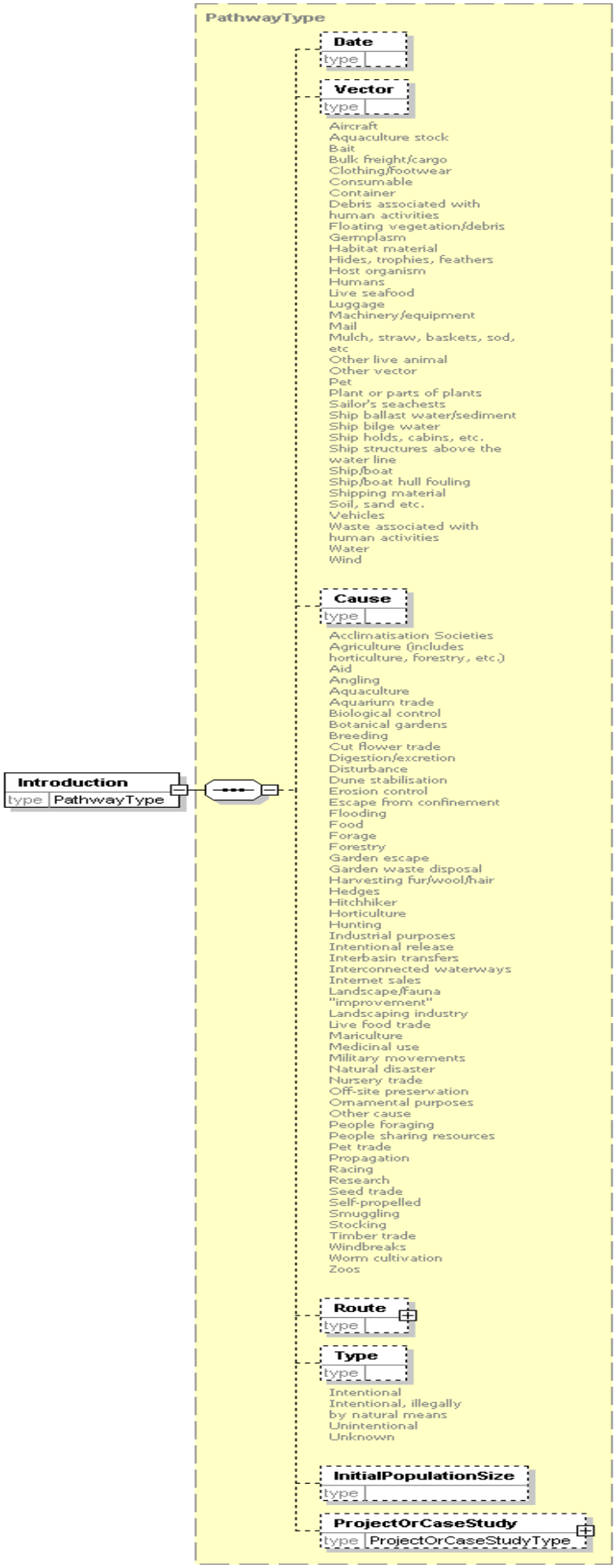
Elements currently available under Biostatus (*Fig. 23*) include Occurrence, Status, Invasiveness, DateOfFirstObservation, Introduction, Dispersal, PopulationDynamics, Distribution, Abundance and Sources. There is the potential to enrich the Occurrence lookup terms, e.g. with additional terms such as incidental, sporadic, ephemerous, adventitious. However terms such as rare, local, common, very common and widespread are best enumerated within the PopulationDynamics (*Fig. 26*), Distribution (*Fig. 27*) and Abundance (*Fig. 87*) elements.

Fig. 24. The Invasiveness element



The Invasiveness element (*Fig. 24*) has a user-defined schema so that users can record the definition of ‘invasive’ that they are using, or they can reference the classification schema they are using; e.g. when they use terms such as ‘potentially invasive, moderately invasive and highly invasive’. The components of the DefinedSchemaType appear in *Fig. 29*.

Fig. 25 The Introduction and Dispersal elements

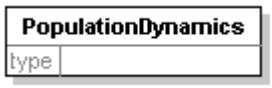


The Introduction and Dispersal elements (*Fig. 25*) have identical internal structures. Users can decide which one is more appropriate depending on the political and spatial context of their data. Introduction information is usually the focus of preventative biosecurity measures whereas Dispersal information is more often the focus of management measures. The Introduction and Dispersal elements contain Date, Vector, Cause, Route, Type (e.g. intentional or unintentional), InitialPopulationSize and ProjectOrCaseStudy elements.

The Cause, Route and Vector structure came from Carlton and Ruiz (2003). 'Cause' is why a species is transported, whether accidentally or deliberately, 'Vector' is how a species is transported, that is, the physical means or agent, and 'Route' is the geographic path over which a species is transported from the origin (donor area) to the destination (target area). The Route element contains To, From, Corridor, and Distance sub-elements.

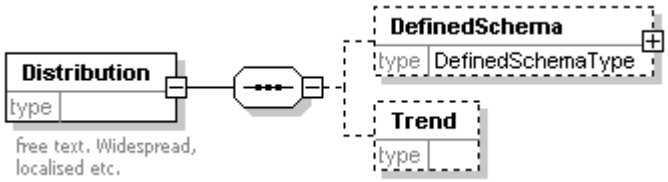
The Introduction and Dispersal elements might be further improved with the inclusion of 3 other elements identified by Carlton and Ruiz (2003). They are VectorTempo; i.e. how a given vector operates through time, in terms of size and rate, speed, and timing, VectorBiota; i.e. description of the biota (the propagules) transferred by a given vector, in terms of diversity, density, and condition, and VectorStrength; the relative number or rate of established invasions that result within a specified time period from a given vector in a particular geographic region.

Fig. 26 The PopulationDynamics element



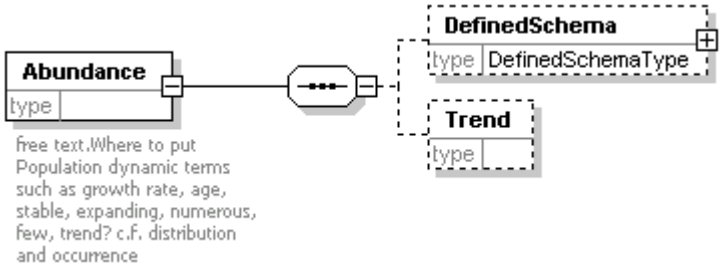
The PopulationDynamics element (*Fig. 26*) is not enumerated in this version of the IAS-PS but it would support information such as "stable/expanding" or "numerous/few", direction of expansion, rate of expansion ("rapid" or "gradual"), size of specimens collected, population growth rate, population age, and initial/current population size.

Fig. 27 The Distribution element



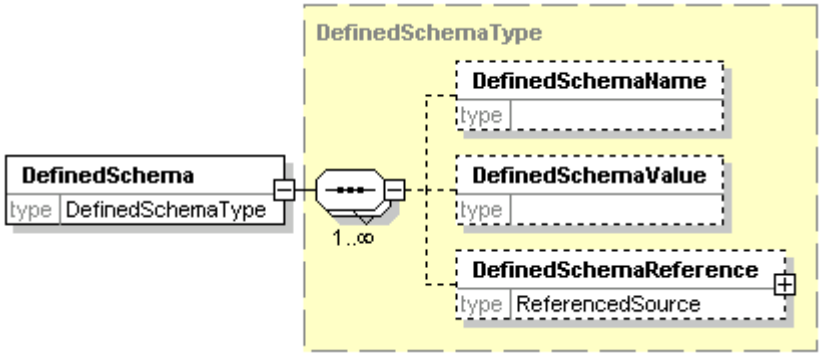
The Distribution element (*Fig. 27*) contains Trend and DefinedSchema sub-elements.

Fig. 28 The Abundance element



The Abundance element (*Fig. 28*) contains Trend and DefinedSchema sub-elements.

Fig. 29 The DefinedSchema element



The DefinedSchema element (*Fig. 29*) is used throughout the IAS-PS to allow users to use and reference their own classification systems.

Appendix 2. List of online databases and other resources examined

Online databases

Agence Méditerranéenne de l'Environnement (AME): Fiches
Alien Invader Plants in South Africa
Alien Species in Poland
Animal diversity Web
Ant Database: Manaaki Whenua-Landcare Research NZ Limited
Applesnail.net
Aquatic Plant Information System Online (APIS)
Baltic Sea Aquatic Species Database
Belgian Forum on Invasive Alien Species
CABI crop, forestry and animal health compendia
Canadian Biodiversity Information Facility
Chesapeake Bay Introduced Species Database
CIESM Atlas of Exotic Species in the Mediterranean Sea
Computer Taxonomy and Ecology of Soil Animals (COMTESA)
Crop Profiles Database (USDA)
Danish Terrestrial Aliens Database
Database on Introductions of Aquatic Species (FAO)
Delivering Alien Invasive Species Inventories for Europe (DAISIE)
Ecological Database of the World´s Insect Pathogens (EDWIP)
European Register of Marine Species

FishBase
Global Compendium of Weeds (GCW)
GloBallast
I3N Invasive Species Database
Introduced Marine Species of Hawaii Guidebook
Introduced Plants and Animals (in Russia)
Invasive Exotic Species in the Iberian Peninsula
Invasive Plant Atlas of New England
Invasive species fact sheet descriptive fields DRAFT (IABIN)
List of Invasive Species of Pakistan
National Aquatic Introduced Species Information System (NIASIS)
National Introduced Marine Pest Information System (NIMPIS)
National PLANTS Database (USDA)
NISbase
Non-Native Aquatic Species in the Gulf of Mexico Region
Nordic Network on Introduced Species (NNIS)
North America Freshwater Fishes Index TNHC
NZ Plant Conservation Network
Pacific Island Ecosystems at Risk (PIER)
Pest Fruit Flies of the World
Plant Viruses online
ScaleNet
Texas Freshwater Fishes Index TNHC
The Brazilian Invasive Alien Species Database
The Nordic/Baltic Network on Invasive Alien Species (NOBANIS)
Universal Virus Database, ICTVdB
USGS Non-indigenous Aquatic Species Database
Viruses of Plants in Australia

Risk assessment tools
Hayes, K. R. and Sliwa, C. (2003). Identifying potential marine pests? A deductive approach applied to Australia. Marine Pollution Bulletin 46: 91-98
http://www.marine.csiro.au/crimp/npl.htm
European Plant Protection Organisation (EPPO) PRA standards
http://www.eppo.org/PUBLICATIONS/pra/pra.htm
Risk assessment for the import and keeping of exotic vertebrates in Australia
http://www.feral.org.au/ref_docs_images/PC12803.pdf

Appendix 3. Commonly used definitions of the word ‘invasive’

IUCN definition:
http://www.iucn.org/themes/ssc/pubs/policy/invasivesEng.htm#anchor392619

"**Alien invasive species**" means an alien species which becomes established in natural or semi-natural ecosystems or habitat, is an agent of change, and threatens native biological diversity.

"**Alien species**" (non-native, non-indigenous, foreign, exotic) means a species, subspecies, or lower taxon occurring outside of its natural range (past or present) and dispersal potential (i.e. outside the range it occupies naturally or could not occupy

without direct or indirect introduction or care by humans) and includes any part, gametes or propagule of such species that might survive and subsequently reproduce.

"**Semi-natural ecosystem**" means an ecosystem which has been altered by human actions, but which retains significant native elements.

CBD definition:
http://www.biodiv.org/decisions/default.asp?lg=0&dec=VI/23
"**Invasive alien species**" means an alien species whose introduction and/or spread threaten biological diversity.

Global Invasive Species Programme (GISP).
http://www.gisp.org/about/IAS.asp
"Invasive alien species" are non-native organisms that cause, or have the potential to cause, harm to the environment, economies, or human health.

Appendix 4. Proposed standard vocabularies

Realm, Habitat and Life form:

Realm lookup	Habitat lookup	Life form lookup
Select one or more	*Select one or more*	*Select one or more*
Terrestrial	Coastland	Virus
Freshwater	Marine habitats	Bacteria
Brackish	Estuaries	Fungus
Marine	Lakes	Arachnid
	Watercourses	Insect
	Riparian zones	Flatworm
	Wetlands	Nematode
	Urban areas	Mollusc (aquatic and terrestrial)
	Agricultural areas	Alga/seaweed
	Disturbed areas	Anemone
	Planted forests	Coral
	Natural forests	Comb jelly
	Scrub/shrublands	Jellyfish
	Range/grasslands	Crustacean
	Tundra	Starfish
	Deserts	Fish
	Ice	Amphibian
	Host	Reptile
	Vector	Bird
		Mammal
		Aquatic plant (floating or submerged)
		Palm
		Tree
		Shrub
		Grass
		Rush
		Sedge
		Herb
		Vine, climber
		Bromeliad
		Fern
		Succulent

Biostatus:

Biostatus		
Occurrence	Status	Invasiveness
Select one	*Select one*	*Select one*
Absent	Alien	Invasive
Recorded in error	Native (no further data)	Not invasive
Extinct	Native. Endemic	Not specified
Eradicated	Native. Non-endemic	Uncertain
Border intercept	Not specified	
Reported	Biostatus uncertain	
Established (able to survive)		
Naturalised (able to survive and reproduce)		
In captivity/cultivated		
Sometimes present		
Present/controlled		
Uncertain		

Impacts:

Impacts		
Impact type	Impact mechanism	Impact outcome
Select one or more	*Select one or more*	*Select one or more*
Environmental	Alleopathic	Altered trophic level
Human health	Causes allergic response	Changed gene pool
Economic / livelihood	Competition-Monopolising resources	Conflict
Cultural / amenity	Competition-Shading	Damaged ecosystem services
	Competition-Smothering	Ecosystem change
	Competition -Strangling	Habitat alteration
	Competition-Other	Host damage
	Disease transmission	Increases vulnerability to invasions
	Filtration	Infrastructure damage
	Fouling	Loss of endangered species
	Herbivory/Grazing/Browsing	Loss of medicinal resources
	Hybridisation	Loss of native species
	Induces hypersensitivity	Modification of fire regime
	Interaction with other invasive species	Modification of hydrology
	Parasitism	Modification of natural benthic communities
	Pathogenic	Modification of nutrient regime
	Poisoning	Modification of successional patterns
	Pollen swamping	Monoculture formation
	Predation	Negatively impacts agriculture
	Rapid growth	Negatively impacts aquaculture
	Rooting	Negatively impacts cultural/traditional practices
	Trampling	Negatively impacts forestry
		Negatively impacts human health
		Negatively impacts livelihoods
		Negatively impacts mariculture
		Negatively impacts tourism
		Obstructs waterways
		Other
		Reduced amenity values
		Reduced native biodiversity
		Selective loss of genotypes
		Soil accretion
		Threat to endangered species
		Threat to native species
		Transportation disruption

Introduction and Dispersal:

Introduction and Dispersal						
Type	**Date**	**Cause**	**Route (geographic path)**			**Vector**
			To	From	Corridor	
Select one	free text	*Cause = enterprise, activity, trade, endeavor, commerce, motive, rationale, incentive, reason* *Of accidental: unintentional, inadvertent, escape, chance. Of deliberate: intentional, planned, purposeful, premeditated, planted, direct.*	free text	free text	free text	*Vector = pathway, mode, dispersal mechanism, transport mechanism, manner, carrier, bearer, method*

Intentional	Acclimatisation Societies	Aircraft
Intentional, illegally	Agriculture (incl. horticulture, forestry, etc.)	Aquaculture stock
Introduced by natural means	Aid	Bait
Unintentional	Angling	Bulk freight/cargo
Unknown	Aquaculture	Clothing/footwear
	Aquarium trade	Consumable
	Biological control	Container
	Botanical gardens	Debris associated with human activities
	Breeding	Floating vegetation/debris
	Cut flower trade	Germplasm
	Digestion/excretion	Habitat material
	Disturbance	Hides, trophies, feathers
	Dune stabilisation	Host organism
	Erosion control	Humans
	Escape from confinement	Live seafood
	Flooding	Luggage
	Food	Machinery/equipment
	Forage	Mail
	Forestry	Mulch, straw, baskets, sod, etc
	Garden escape	Other live animal
	Garden waste disposal	Other vector
	Harvesting fur/wool/hair	Pet
	Hedges	Plant or parts of plants
	Hitchhiker	Sailor's seachests
	Horticulture	Ship ballast water/sediment
	Hunting	Ship bilge water
	Industrial purposes	Ship holds, cabins, etc.
	Intentional release	Ship structures above the water line
	Interbasin transfers	Ship/boat
	Interconnected waterways	Ship/boat hull fouling
	Internet sales	Shipping material
	Landscape/fauna "improvement"	Soil, sand etc.
	Landscaping industry	Vehicles
	Live food trade	Waste associated with human activities
	Mariculture	Water
	Medicinal use	Wind
	Military movements	
	Natural disaster	
	Nursery trade	
	Off-site preservation	
	Ornamental purposes	
	Other cause	
	People foraging	
	People sharing resources	
	Pet trade	
	Propagation	
	Racing	
	Research	
	Seed trade	
	Self-propelled	
	Smuggling	
	Stocking	
	Timber trade	
	Windbreaks	
	Worm cultivation	
	Zoos	

REFERENCES:

Carlton, J.T., and Ruiz, G.M., 2003. Principles of Vector Science and Integrated Vector Management, in: H. Mooney et al., Editors, Best Practices for the Prevention and Management of Alien Invasive Species. Island Press.

Colautti, R.I. & MacIsaac, H.J. (2004) A neutral terminology to define 'invasive' species. Diversity and Distributions, 10, 135–141.

Hilliard, R., 2004. Best practice for the management of introduced marine pests: a review. Prepared for GISP by URS Australia Pty. Ltd.

Richardson, D., Pysek, P., Rejmánek, M., Barbour, M., Panetta, F.D., and West, C. 2000. Naturalization and invasion of alien plants: concepts and definitions. Diversity and Distributions (2000) 6, 93–107