Biodiversity Information Standards
TDWG

Mechanisms for coordination and delivery of taxon profiles in Australia

Longitudinal use case scenarios from primary data custodians and roles for data standards

Alex R. Chapman

Department of Environment and Conservation
Western Australian Herbarium

Structured taxonomic descriptive data standards

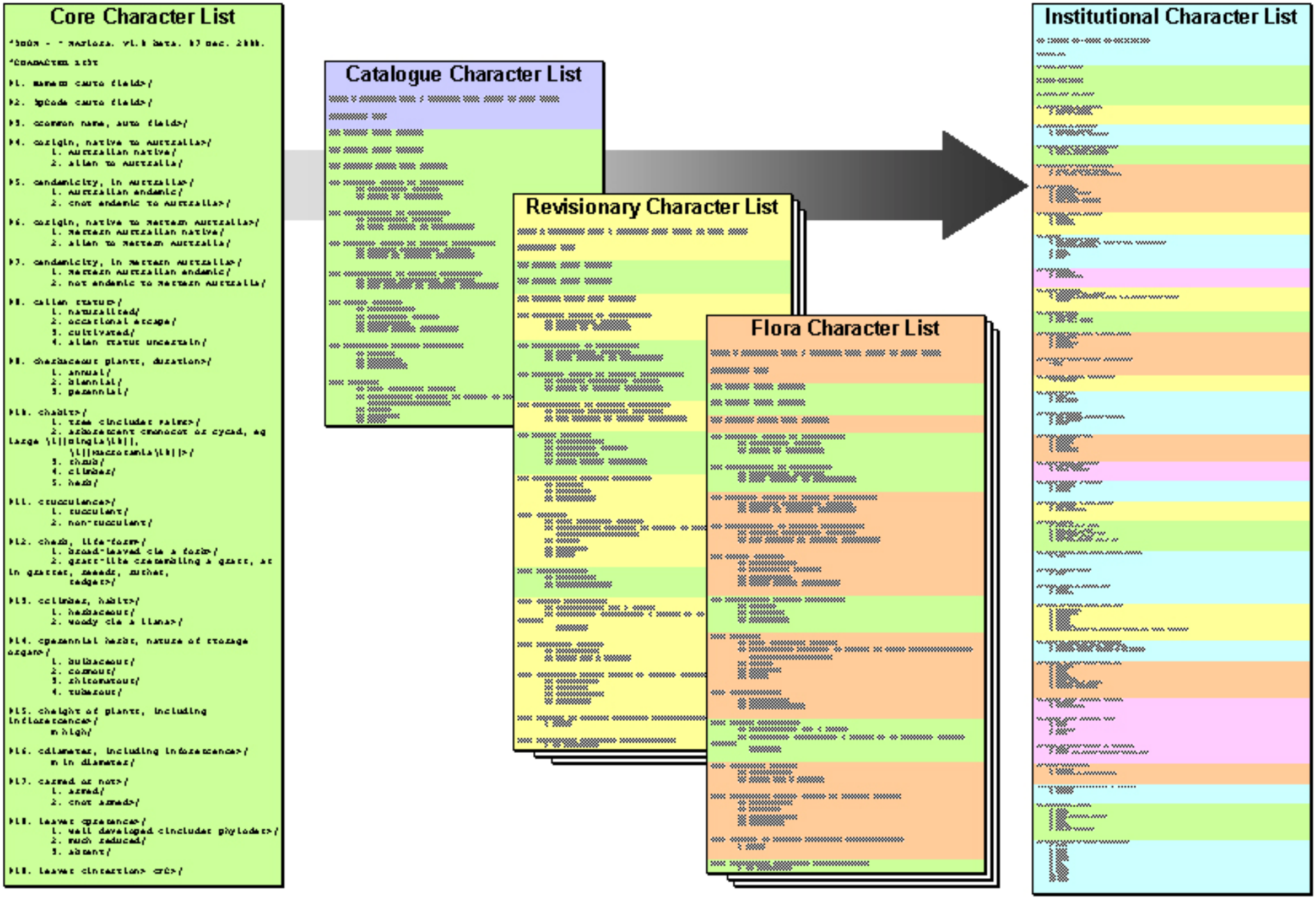
- enable the rigorous capture of taxon data
- at an atomic character level
- for use in -
  - information retrieval
  - interactive identification
  - rendering descriptions
    esp. within taxon profiles
  - data harvesting & aggregation

Coordinating the capture of coded descriptive data

Requires an agreed vocabulary & the definition of characters and states

- for one taxonomic group
  - becomes more complex with additional contributors
- across taxonomic groups
  - becomes more complex with increased variety of taxa and characters

From 1992 the WA Herbarium developed a model for coordination


Core Character List
Catalogue Character List
Revisionary Character List
Flora Character List
Institutional Character List

Progress

- Developed the backend tools to manage multiple datasets

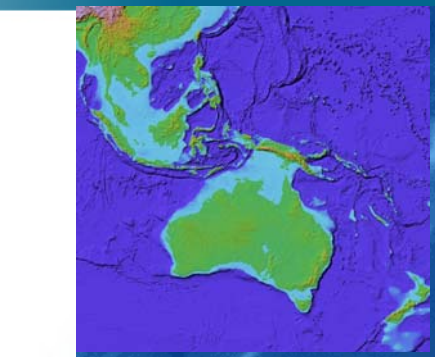
  Uptake slow due to the:

  - Lack of frontend tools, and workbench technologies to support specialist needs
  - Strict requirements of methods
  - Large learning curve
  - Lack of interest in specialists to build collaboratively, or
  - No pay-off for spending time contributing to non-paper publications

Mechanisms for coordination and delivery of taxon profiles in Australia

Western Australia

- 2,500,000 km^2
  -- *c.* ⅓ of continent
- over 13,000 vascular species
  -- *c.* ½ the flora
- 3 bio-provinces
- 79% endemism in the mediterranean SW



Biodiversity Information Standards
TDWG

Mechanisms for coordination and delivery of taxon profiles in Australia

FloraBase
the Western Australian Flora

Advanced
Name Quick Search:
Go

Home | Themes | Find | Tools | Help

Plant Names

Flora Descriptions

Interactive Identification

Species Images

Specimen Data

Distribution Maps

Botanical Library

Plant of the Month

Latest Projects

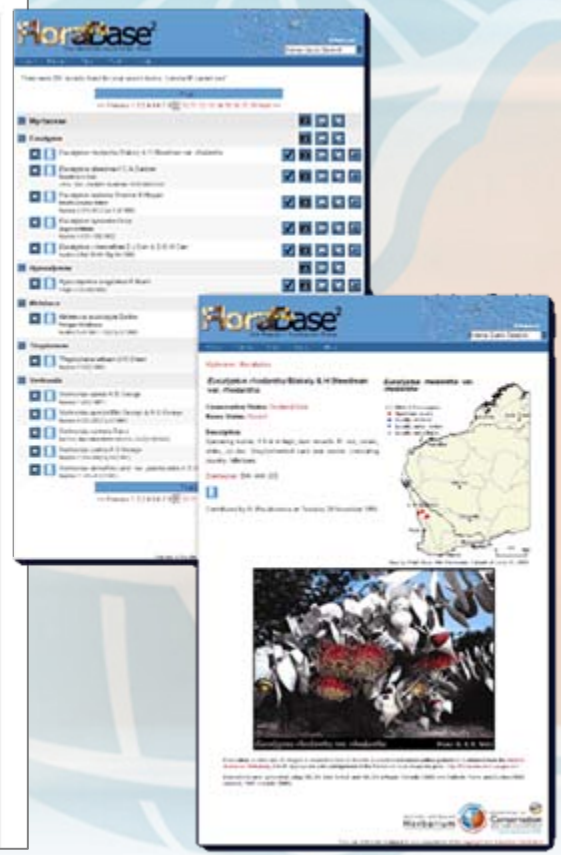
News

Minister launches Interactive Keys
14 Sep 2007
The Minister for Environment and Climate Change launches our latest innovation — online identification tools — in the new FloraBase 2.5.

September PotM: *Dampiera lavandulacea*
1 Sep 2007
This small shrub or perennial herb, with its flowers of deep blue and sometimes purple, is a stand-out feature in the bushland from July to November.

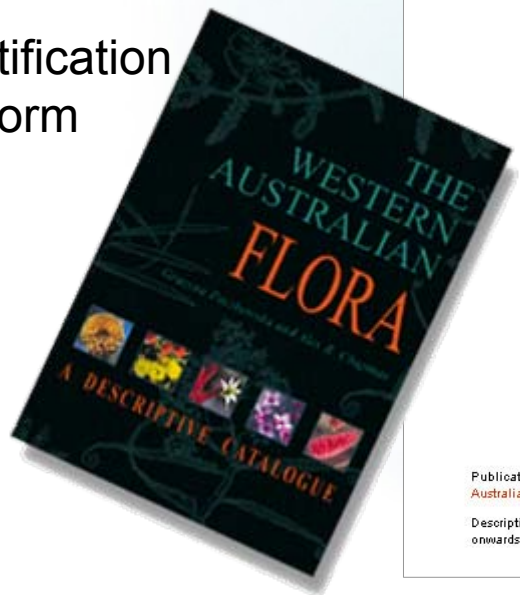
New Churchill Fellow: Ray Cranfield
15 Aug 2007
To study historical lichen collections and community participation in lichen surveys and collections - UK, Sweden.

2007 flora stats published
8 Jun 2007
The latest authoritative stats on species names in use in WA, and a baseline count of the number of cryptogams tracked in the State Census.



Descriptive Data

- DELTA-coded data for
  13,000 species x 10 characters
- 692-page book published in Oct. 2000
- Data provides interactive identification via web query form in FloraBase



FloraBase² the Western Australian Flora

Advanced
Name Quick Search:

Home Themes Find Tools Help

Myrtaceae : Eucalyptus

Eucalyptus rhodantha* Blakely & H.Steedman var. *rhodantha

Conservation Status: Declared Rare
Name Status: Current

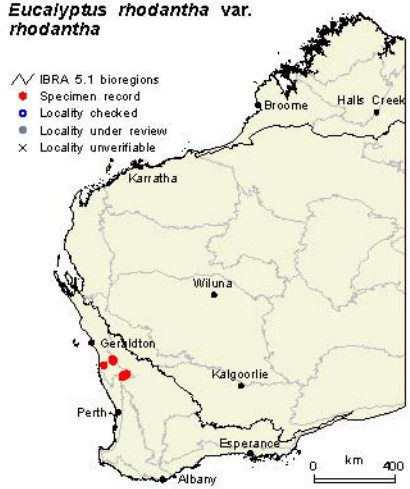
Description:
Spreading mallee, 1.5–4 m high, bark smooth. Fl. red, cream, white, Jul–Jan. Grey/yellow/red sand over laterite. Undulating country, hillslopes.

Distribution: SW: AW, GS.

R

Contributed by G. Paczkowska on Tuesday 28 November 1995

Eucalyptus rhodantha* var. *rhodantha

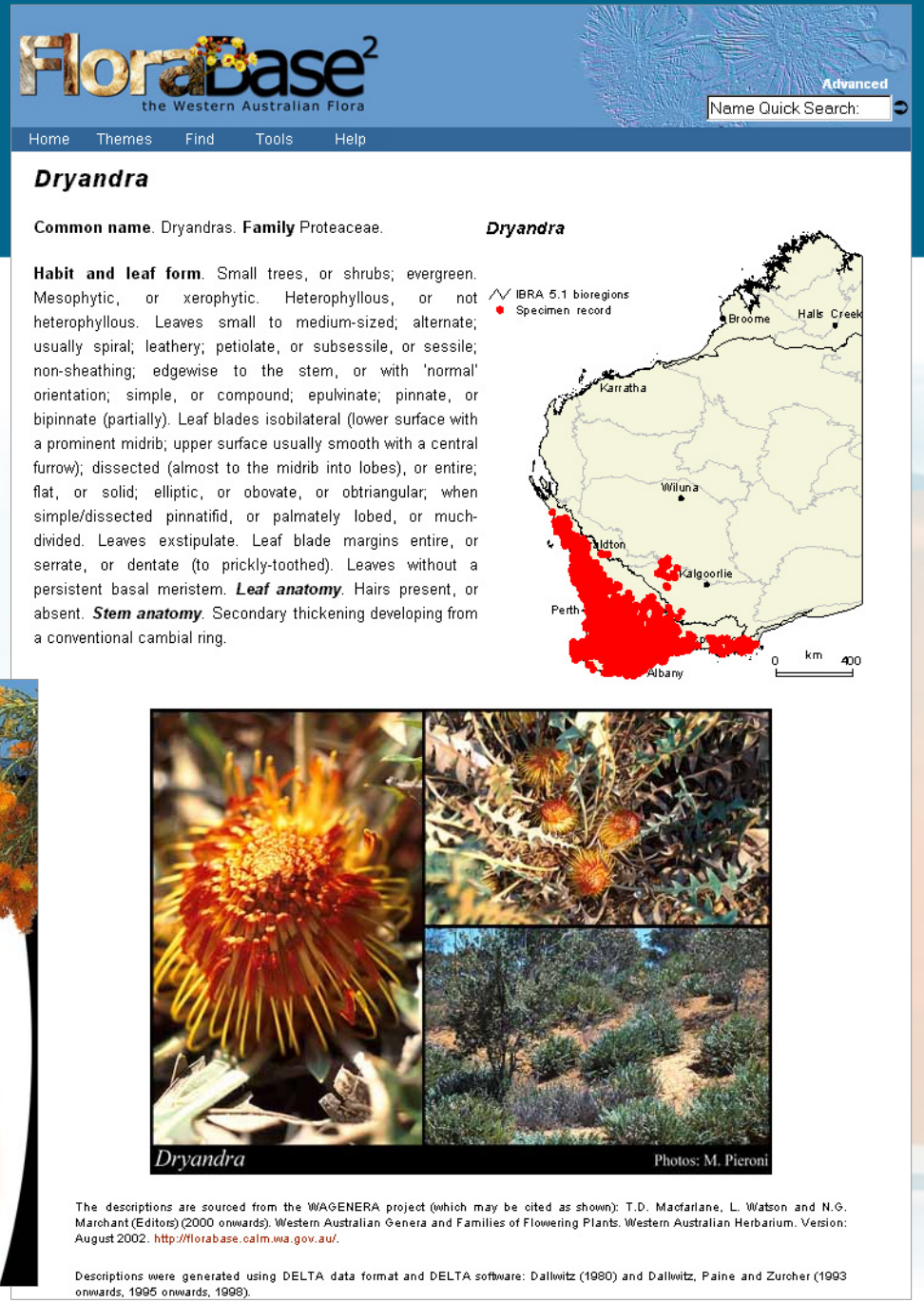


Map by Paul Gioia, WA Herbarium. Current at June 13, 2003

Descriptions were generated using DELTA data format and DELTA software: Dallwitz (1980) and Dallwitz, Paine and Zurcher (1993 onwards, 1995 onwards, 1998).

Descriptive Data

- DELTA-coded data for
  1,200 genera, 200 families x 800 characters
- Interactive keys in Sep. 2007



Other Western Australian information products

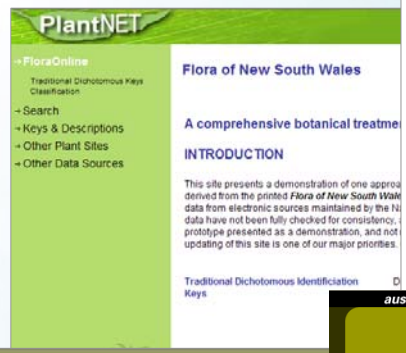
- Flora descriptions
- Protologues and other scientific taxonomic descriptions
- Interactive key content
- Other databases

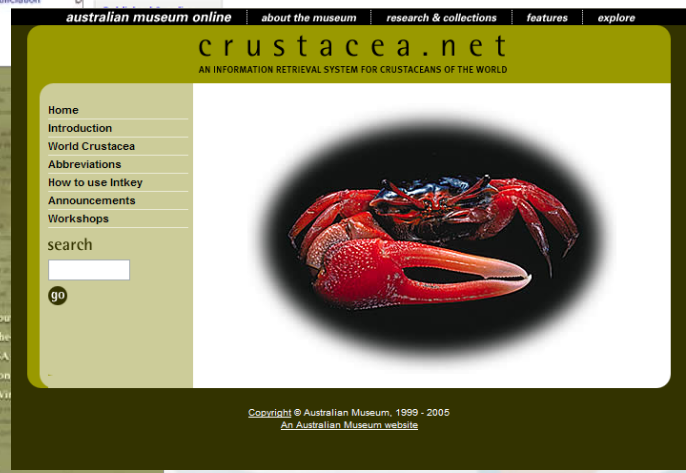
ie. publication model now more open to partially-structured descriptive text
(but still need to manage it)

Other Australian information products

- Flora descriptions
- Protologues and other scientific taxonomic descriptions
- Interactive key content
- Databases
- Web portals

nb. many web portals often only package static descriptions





Scalability aspects

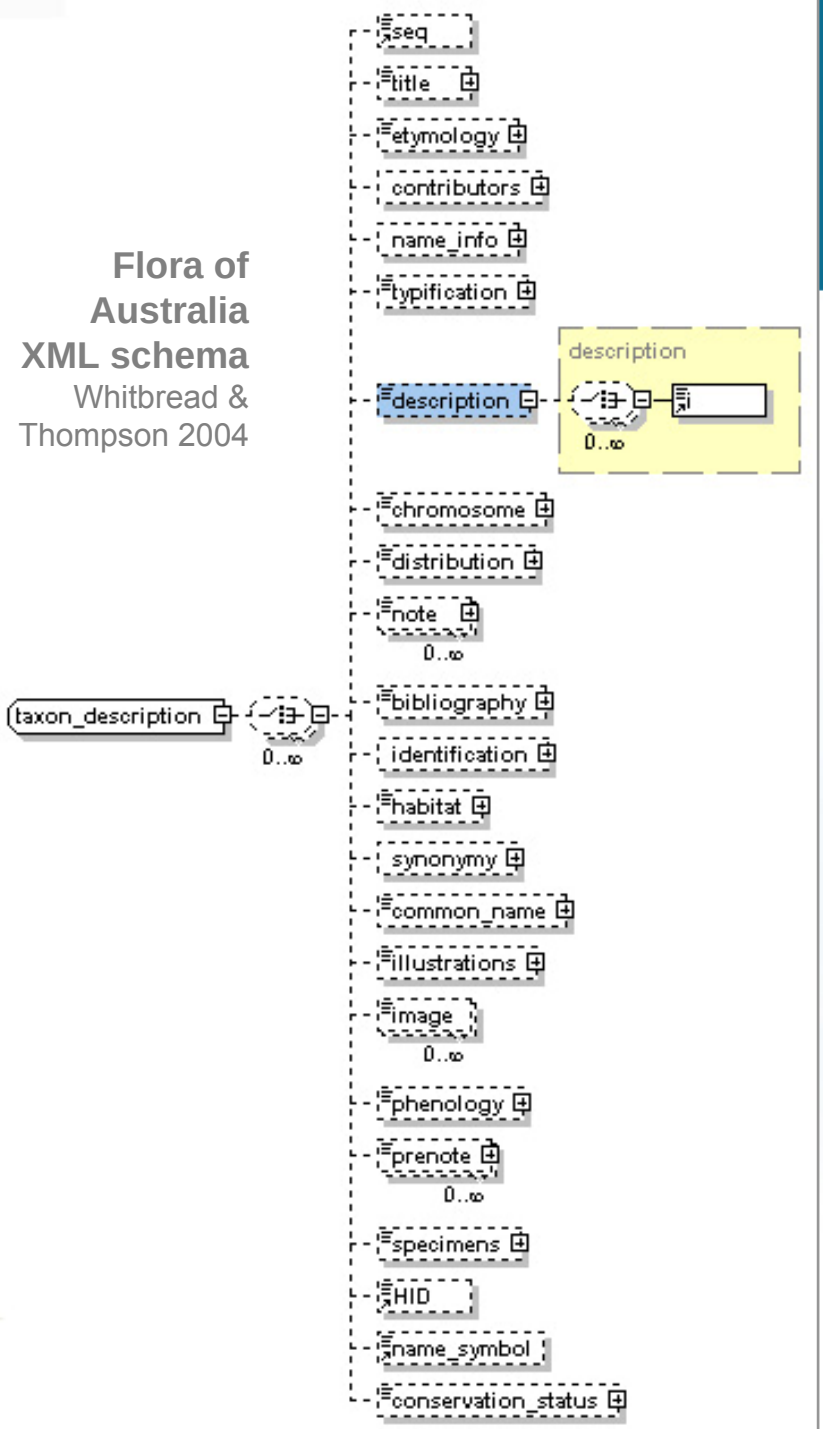
- *Flora of Australia* volumes:
  - are marked up with XML schema identifying blocks of descriptive text
  - may provide a common reference standard for marking up species-level descriptions from the many potential sources

- The *Atlas of Living Australia* and *Encyclopaedia of Life* projects will require:
  - the identification of standard components on taxon profile pages
  - a mechanism for reliable tagging

- TDWG Species Profile Model:
  - an RDF schema that identifies many of the same objects
  - An umbrella schema for delivering structured (SDD) & unstructured data

Mechanisms for coordination and delivery of taxon profiles in Australia

Flora of Australia XML schema
Whitbread & Thompson 2004



Species Profile Model (simplified RDF schema)
Hyam & Doering 2007

▼ **Class**		
Species Profile Model	5 predicates	5 property values
Conservation	6 predicates	6 property values
Physiology	6 predicates	6 property values
Size	6 predicates	6 property values
Evolution Info Item	6 predicates	6 property values
MolecularBiology	6 predicates	6 property values
Use	6 predicates	6 property values
Cytology Info Item	6 predicates	6 property values
Biology	6 predicates	6 property values
Information Item	5 predicates	5 property values
Distribution	6 predicates	6 property values
Ecology	6 predicates	6 property values
Description	6 predicates	6 property values
▼ **ObjectProperty**		
About Taxon	6 predicates	6 property values
Context Occurrence	6 predicates	6 property values
Context Value	6 predicates	6 property values
Has Content	6 predicates	6 property values
Has Information	6 predicates	6 property values
7 more...		
▼ **Ontology**		
TDWG SpeciesProfileModel LSID Ontology	8 predicates	8 property values
▼ **DatatypeProperty**		
Has Content	6 predicates	6 property values
context	6 predicates	6 property values

Challenges for:

- Primary data providers, custodians and publishers:
  - finding a way to meet their local aims and needs while contributing their content upwards to national and international initiatives
  - while retaining credit and recognition for their efforts
- Aggregators:
  - harvesting responsibly, maintaining currency of their views, facilitating collaboration, recognising data providers
- BIS(TDWG):
  - providing clear implementable standards and
  - upgrade pathways for legacy standard datasets
- Consumers:
  - discriminating content quality and reliable sources

2008

Biodiversity Information Standards

TDWG

19- 25 October