

# METADATA

Data about data

GSN Indibano

March 2010

Port Elizabeth



# Metadata

- Metadata are defined as the ‘data about data’
- The purpose of metadata is to describe data so that it can be used in the future with no confusion about its content or purpose
- Metadata can be created for a vast range of electronic and physical items such as raw data, books, CDs, DVDs, images, maps, database tables and web pages
- A number of different fields can be captured and constitute metadata



# Metadata Standards

- Metadata elements grouped into sets and designed for specific purposes are called *metadata schemes*
- For every metadata element, there are rules governing the content of that element, representation and allowed content
- Many schemes use XML to specify their syntax and many are developed and maintain by standard organizations such as ISO
- Many different metadata standards exist and are applied to specific types of data



# Metadata Standards

Standard	Application
Darwin Core	Geographic occurrence of species and the existence of specimens within a collection
Dublin Core	General Purpose
EML	Ecological data - species
ISO19115	Geographic information
SANS1878	Spatial data
FGDC	Geographic data

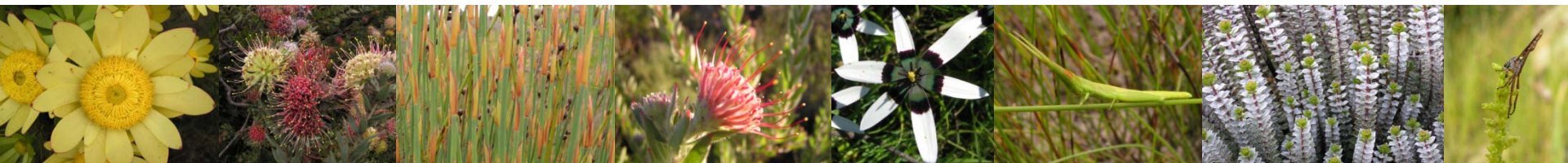
This is only a tiny fraction of all the different metadata standards





# Typical Metadata fields

Title	Subject	Date	Creator
Individual	Organization	Address	Contributor
Keywords	Abstract	License	Usage Rights
Methodology	Language	Format	Object Name
Attribute Name	Definition	Type of Value	Units
Location	Longitude	Latitude	Start Date
End Date	Temperature	Taxonomy	Online distribution
Metadata standard	Custodian	Contact Information	Sampling technique



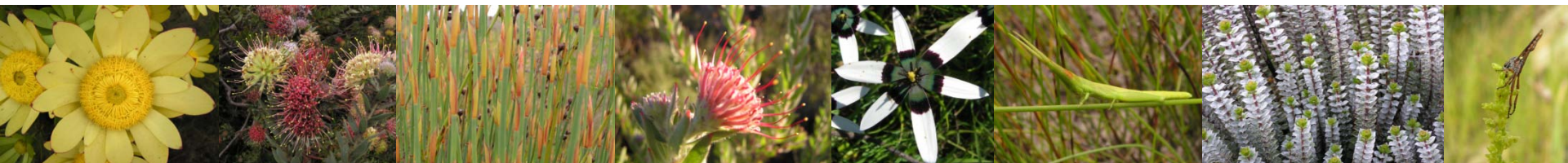
# XML Document

```
<?xml version="1.0" ?>
<eml:eml xmlns:eml="eml://ecoinformatics.org/eml-2.0.1" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" packageId="nikkis.43.15" system="knb"
  xsi:schemaLocation="eml://ecoinformatics.org/eml-2.0.1 eml.xsd" scope="system">
- <dataset scope="document">
  <title>Nest mounds of the granivorous harvester ant Messor capensis, Tierberg, 1991 - 2005</title>
- <creator id="1203343259718" scope="document">
  - <individualName>
    <salutation>Dr</salutation>
    <givenName>W. Richard J</givenName>
    <surName>Dean</surName>
  </individualName>
  <organizationName>Percy FitzPatrick Institute of African Ornithology, University of Cape Town (now retired).</organizationName>
- <address scope="document">
  <deliveryPoint>8 Nuwe Street</deliveryPoint>
  <city>Prince Albert</city>
  <postalCode>6930</postalCode>
  <country>South Africa</country>
</address>
  <phone phonetype="voice">023 541 1828</phone>
  <electronicMailAddress>Richard Dean [lycium@telkomsa.net]</electronicMailAddress>
</creator>
- <associatedParty id="1203343643937" scope="document">
  - <individualName>
    <salutation>Dr</salutation>
    <givenName>Suzanne Jane</givenName>
    <surName>Milton</surName>
  </individualName>
  <organizationName>Percy FitzPatrick Institute of African Ornithology, University of Cape Town (now retired).</organizationName>
- <address scope="document">
  <deliveryPoint>8 Nuwe Street</deliveryPoint>
  <city>Prince Albert</city>
  <postalCode>6930</postalCode>
  <country>South Africa</country>
</address>
  <phone phonetype="voice">023 541 1828</phone>
  <role>Collaborator</role>
</associatedParty>
- <abstract>
```



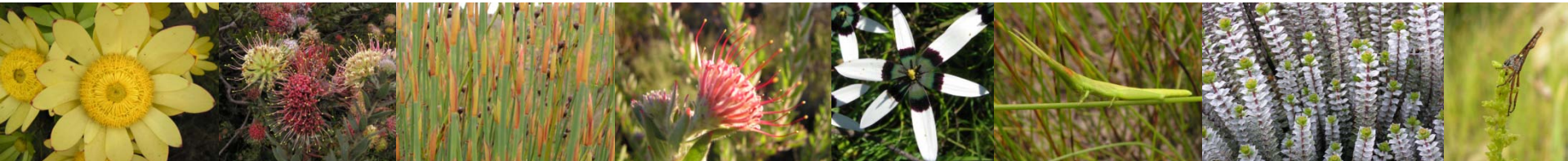
# Metadata Creation

- Various software programs exist which make the creation of metadata compliant with the relevant standards
- Examples: Morpho / GeoNetwork
- HOWEVER – you do not need fancy software, you can create metadata about any dataset, in any format that you want!
- Record descriptive fields about your data and store them in the data file – otherwise you WON'T remember what you did



# Morpho

- Morpho is used to create EML metadata
- [Morpho.1.7.0.](#)





# Importance

- It needs to be **good** metadata – rubbish in → rubbish out
- It helps YOU remember how you collected your data, what the units were, what the sampling methods were etc.
- It helps you remember what data you have, what is still outstanding, potential integration across datasets and what potential analyses you can run
- Data archiving – good metadata allows you to find the dataset and find something that is worthwhile for your research
- Other people can find YOUR data, and your work can be re-used and contribute to science

