



**KING'S**  
*College*  
**LONDON**

---

---

## **The DTC Data Archive and what we will get out of it**

---

Mike Dobson: *Freshwater Biological Association*

# Purpose of the archive

---

- Curating data generated and captured by research projects
- Researchers create data, we have to make it useful.
- Data archive, but also querying, browsing, visualising, analysing, other interactions
- Need to meet needs of different users – researchers, also land managers, civil servants, planners, CSF Officers, etc.

# The Data

---

- Mostly numerical in some form: spreadsheets, databases, CSV files
  - Sensor data (automated, telemetry)
  - Manual samples/analyses
- Species/ecological data
- Geo-data
- Also less highly structured information:
  - Time series images, video
  - Stakeholder surveys
  - Unstructured documents

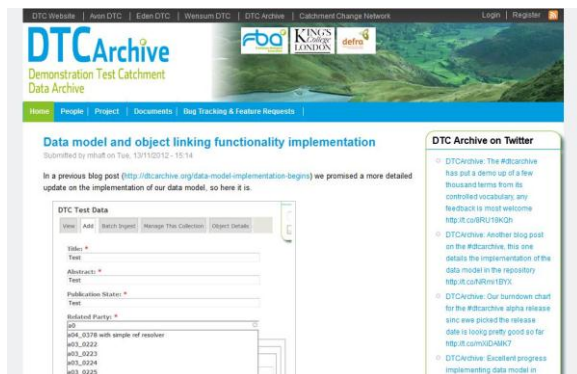
# Why use the Archive?

---

- Defra funded research must comply with the INSPIRE Directive
  - Includes standards for describing, representing, disseminating geo-spatial data.
  - Standards compliant data is much more re-useable by researchers
- Data held in the archive will be stored in an appropriate format for long term digital curation
- Data in the Archive will be given a DOI and so will be published and citable

# How will people access data?

- Need to consider:
  - different types of user
  - different points of access: e.g. Defra website, DTC catchment websites, DTC Archive website, search engine



# Making datasets useful

---

- Needs to use a controlled vocabulary:
  - a **set of agreed terms**, each with an unambiguous definition.
- **Creation requires work**, but it allows **greater accessibility** and **documentation** of relevant datasets.
- i.e. it makes sharing, understanding and searching easier.

---

For example, four scientists write their outputs and label the phenomena measured as:

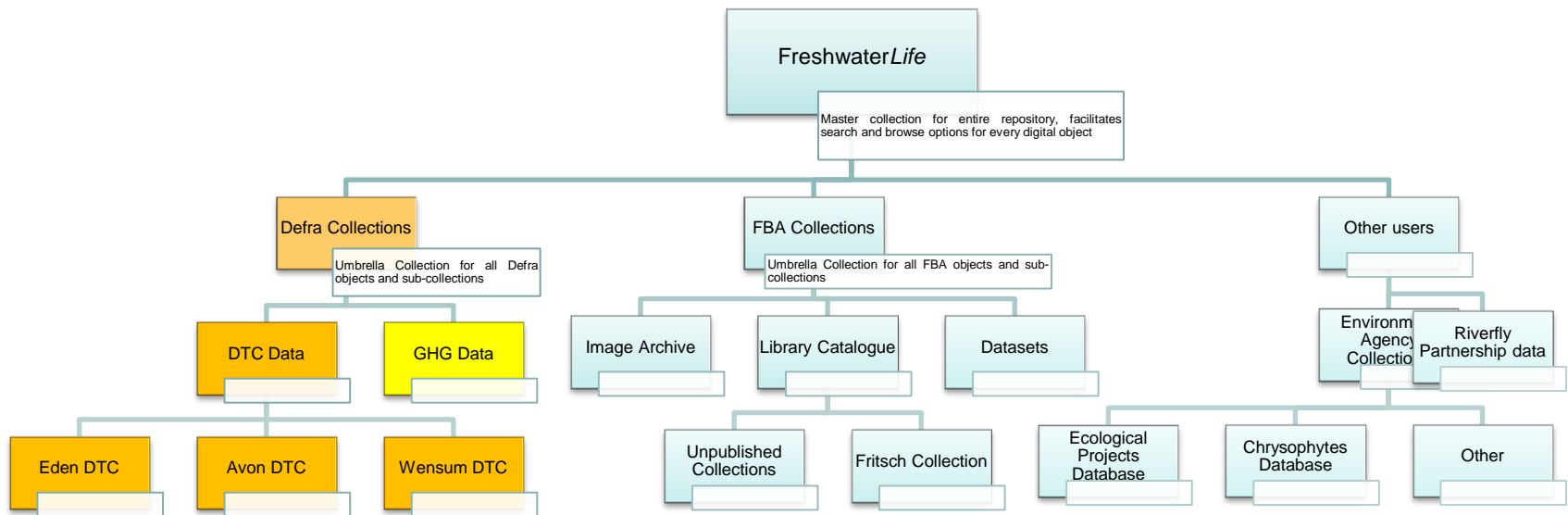
- T (degC) - Temperature (in degrees C)
- T - Time of day (HH:MM:SS)
- Temp - Temperature (in Kelvin)
- Tmp - Trimethyl phosphate volume (ml)

Agreeing standard definitions and identifiers/labels removes this ambiguity.

(Thanks to Ag Stephens, CEDA)

# Interconnections among data sources

## FBA Collections structure



Searches in one may pick up items of relevance in others



- 
- Interoperability via international standards will allow this to occur among different collections:
    - DTC Archive and AGHG Archive
    - Defra archives and other FBA data holdings
    - External sources:
      - NERC Virtual Observatory
      - NERC Environmental Information Data Centre holdings
      - National Biodiversity Network

# Thank you

---

General enquiries: [director@fba.org.uk](mailto:director@fba.org.uk)

Technical enquiries: [mhaft@fba.org.uk](mailto:mhaft@fba.org.uk)

<http://dtcarchive.org/>