ANNEX 3: WGZE METADATA GUIDELINES FOR (ZOO) PLANKTON DATA

DRAFT, 24 February 2003

What is plankton metadata and why is it important?

Plankton "metadata" is ancillary information about and related to the plankton data itself, such as the methods and processes involved in measuring or observing these "data" and the conditions under which it was sampled. For zooplankton data, this can be key information, such as net mesh size or a specific sample processing protocol, which specifies the target species and relative content of that data. This additional information is needed to examine and utilise the plankton data in a meaningful and appropriate way. While this information was available at the time of the sample collection, present in a log book or cruise report, special effort must be made to ensure that this metadata remains with the data indefinitely, preferably in a digital form stored with the data itself. By doing this, the quality and usability of the plankton data are preserved for present and future application and study.

What is the purpose of this document?

The Working Group on Zooplankton Ecology (WGZE), with guidance from the Working Group on Marine Data Management (WGMDM), is providing these general metadata guidelines for plankton data collected and submitted to ICES. The existence of such guidelines will ensure that quality and usable plankton data sets will be preserved and available to ICES in the present and future. The metadata fields in this document do not cover every possible metadata type, nor will every plankton data type include all of these metadata fields. Instead, these fields are intended to serve as an example of the types of information (metadata) important towards preserving the highest level of quality and understanding within the zooplankton data. It should also be noted that while these guidelines were written for zooplankton data, they are also highly appropriate for other plankton types (*e.g.*, phytoplankton and bacteria. These may tend to have less emphasis on gear type and greater emphasis on the microscope and/or staining techniques used in processing them).

(ZOO) PLANKTON DATA GUIDELINES:

The fields below represent many of the most common metadata fields important to preserving high quality plankton data. These fields are not necessarily complete nor is every field mandatory. Many of these data-fields may be better stored along with each station or tow within the plankton data sheet itself. The intent is to describe the general types of information that should be preserved either within the data or as a separate metadata or cruise summary.

Metadata-fields relating to the entire CRUISE:

This is general information relating to the sampling cruise. This information may help link these data to physical and chemical measurements taken on the same cruise (but stored in a separate data file or location), and can also give credit to the investigators participating in the cruise.

- Name of the ship
- Investigator-designated Cruise Identifier (e.g., "9801 Leg 1", or "Froggy Fjord 1993")
- Associated Project
- Associated Institute
- Principal investigator(s) for cruise
- Other responsible investigators, and their variable(s)
- If a cruise or data report is available describing the data collection and processing, this can be referenced or, when possible, supplied with the data.

Metadata-fields relating to a specific STATION:

This is specific information relating to the position and time of the sampling station, along with the weather conditions and other details observed during the sampling.

- **Station latitude and longitude** (noting any hemisphere indicators such a "N" for North or negative (-) for South, etc.)
- **Station Month, Day, Year**
- **Station Time** (designated as "local", "GMT/UTC", "ship", etc)
- Investigator-designated **Station Identifier** (e.g., "Station 1x", "Station 2x", ...)
- Optional general station time (twilight, midnight, day, morning)
- **Meteorological Observations** (windy, wavy, cloudy, sunny)
- **Station Sounding** (bottom depth)
- Information about any other supplementary/complementary data collected at the same time (same station) should also be supplied (i.e., a note that "CTD and nutrient samples were also made at this station")
- **Note any affecting instances or corrections applied** (e.g., "a substantial phytoplankton bloom was present at this station" or "a larger net mesh size was used at this station due to frequent clogging by gelatinous zooplankton")

Metadata-fields relating to the NET TOW or BOTTLE CAST:

This information describes the towing (or bottle deployment) methods and procedures.

- **Towing Method** (horizontal, vertical, oblique)
- **Towing depth-range** (a range of starting and ending depths for each net or bottle), or the wire angle and wire out during the tow
- Towing Duration (minutes or hours)
- Towing Distance (in metres)
- Average Towing Speed (knots or metres per second)
- **Note any affecting instances or corrections applied** (e.g., "the gear hit the bottom midway through the tow")

Metadata-fields relating to the SAMPLING GEAR:

This information describes the sampling gear employed, with key metadata fields such as the effective mesh size of the sampler.

- **Describe the sampling gear used**, providing a literature reference if available
- If using a "standard" net (e.g., a NORPAC net) was used, be sure to note any modifications to this net
- **What net mesh size was used** (usually in microns)
- What was the net opening shape (square or circular) and the opening mouth area or diameter
- **Was a flowmeter used?** When and how was is calibrated?
- **Note any affecting instances or corrections applied** (e.g., did the flowmeter break or the codend crack)

Metadata-fields relating to SAMPLE PROCESSING:

This information describes the sample processing methods and protocols.

- **What volume of water was filtered to yield this sample** (i.e., from the flowmeter or calculated via mouth area and towing distance, or estimated because the flowmeter failed)
- **How were samples preserved**, and in what (e.g., 5% buffered formalin)
- **How were samples processed** (summarize the counting, weight, or volume method)?
- **Was the sampled split** (via Folsom splitter or other method)? What was the size of the final aliquot?
- **Were large plankters removed** prior to making biomass measurements? Was a size or volume criteria used in deciding what to remove and what could remain?

- Investigator-designated tow, net, or sample identifier (e.g., “Sample 1035 from Net 5”)
- **Note any affecting instances or corrections applied** (e.g. “eggs and fragments were not counted”)

Metadata-fields relating to SAMPLE ITSELF:

This measured plankton data fields should be clear enough to be understood by others with slightly less expertise and situational knowledge than the original investigator. The tow most common mistakes are not providing units for each measurement (*e.g.,* “number per cubic meter” or “milligrams wet mass per sample”), and not providing clear column headings for the data (*e.g.,* what is “CfcV” and “HetBact”?)

- Provide the units for each measurement (*e.g.,* #/liter, #/m3, #/m2, mg/m3, mg/haul, ...)
- If taxonomic codes, symbols, or abbreviations are used in the data, provide a translation table to help reduce possible misunderstandings of the taxa (e.g., “CfcV” = *“Calanus finmarchius* copepodite V*”, ...)*
- Is an estimate of final uncertainty of the data known? (This is a JGOFS thing, I am not sure if it applies to plankton data.)

Additional formatting and metadata suggestions are available through the Formatting Guidelines for Oceanographic Data Exchange (http://www.ices.dk/ocean/formats/getade_guide.htm) prepared by the IOC's Group of Experts on the Technical Aspects of Data Exchange (GETADE).

An example metadata summary:

Here is a brief metadata summary based on a zooplankton present in the JGOFS AESOPS online data system. Note that without looking at Dagg (1993), it does not appear that a net mesh size is provided.

Principal Investigators: Michael Dagg and Juanita Urban-Rich

Project/Study: US JGOFS Antarctic Environments Southern Ocean Process Study (AESOPS)
Antarctic Polar Front Zone (APFZ) Process 1 and 2 cruises aboard R/V Roger Revelle cruises 7 (Process 1) and 9 (Process 2)

Cruise: “Kiwi-7”

Sampling Gear: Bongo Net

Sampling Procedure: Vertical bongo nets were taken in through the upper 200 m. The general sampling procedure is described in Dagg (1993). The codend contents were preserved with 10% buffered formaldehyde. Replicate aliquots of the sample were counted in the laboratory such that 30 of the dominant large copepod species were counted and greater than 1000 zooplankton per sample counted. An aliquot of the cod-end contents was filtered onto 153 micron Nitex and frozen in liquid nitrogen at sea for dry weight analysis in the laboratory. Copepods were sorted by species, rinsed, dried and weighed on a Cahn microbalance for dry weights.

Dagg, M. J. 1993. Grazing by the copepod community does not control phytoplankton production in the subarctic Pacific Ocean. Progress in Oceanography, 32, 163-183.