Standardizing Metadata for Environmental Microplastics Research

Rodney D. L. Smith
Assistant Professor, Chemistry
Sept. 27, 2021

UNIVERSITY OF WATERLOO

Learning From the Past and Looking Forward

A situation to avoid

Challenges Remain in Combining Data from Multiple Organizations

Since 1899, nutrient records collected by **488** agencies / organizations *at* **321,927** stream sites

for a total of **25** million nutrient records

14.5 million of these nutrient records had missing or ambiguous information that limited their use for regional or national assessments

"At current costs, the nutrient records with **missing or ambiguous reference information represent an estimated $12 billion worth of data** that are unavailable for regional or national analyses by secondary data users."

US Geological Survey
https://acwi.gov/nawqa/NLCmeetings/multisource_data.pdf
Sprague, Oelsner, Argue *Water Research* **2017**, *110*, 252-261

Coming soon to a lab near you

Government of Canada | Gouvernement du Canada

Search Canada.ca

MENU

Home > Interagency research funding > Policies and Guidelines > Research Data Management

Research Data Management

Tri-Agency Statement of Principles on Digital Data

Tri-Agency Research Data Management Policy

4. Implementation dates

...

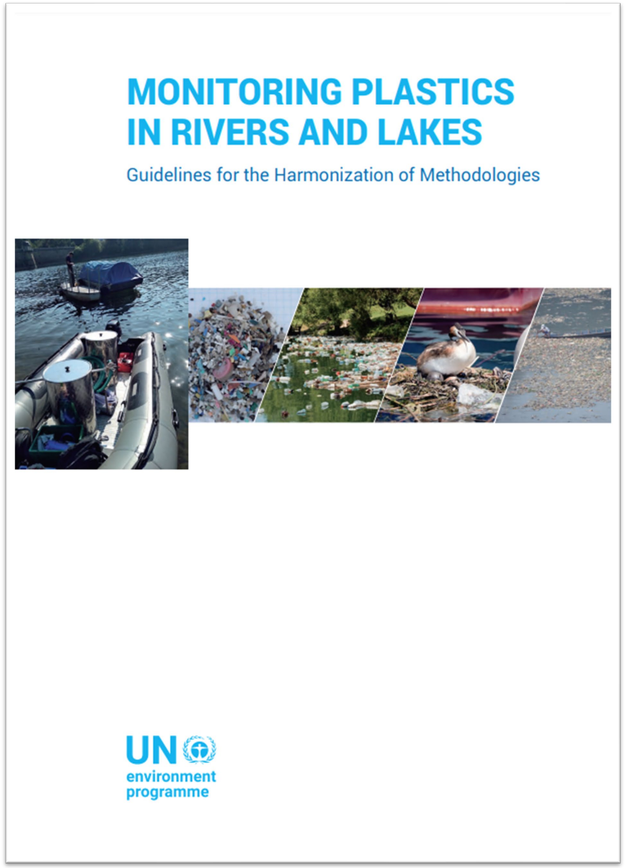
- *Institutional strategies:* **By March 1, 2023**, research institutions subject to this requirement must post their RDM strategies and notify the agencies when they have been completed.
- *Data management plans:* **By spring 2022**, the agencies will identify the initial set of funding opportunities subject to the DMP requirement. The agencies will pilot the DMP requirement in targeted funding opportunities before this date.
- *Data deposit:* After reviewing the institutional strategies and in line with the readiness of the Canadian research community, the agencies will phase in the deposit requirement.

https://www.science.gc.ca/eic/site/063.nsf/eng/h_97610.html

Overarching Objective

Create metadata template for general use in environmental (water and sediments) microplastics research projects.

- adapt existing metadata schemas
- engaging microplastics researchers, stakeholders and RDM practitioners
- establish an ongoing dialogue to adapt to the inevitable evolution of microplastics research



"**To minimize barriers to data exchange**, the format of data and metadata should be compatible with **national or international database structures**. In particular, the **global water quality database GEMStat** (https://gemstat.org) can serve as an example. It should be noted that plastic is currently not in the GEMStat parameter list, although its inclusion is under consideration."

Monitoring Plasics in Rivers and Lakes: Guidelines for the Harmonization of Methodologies
UN Environment Programme (2020)
https://wedocs.unep.org/bitstream/handle/20.500.11822/35405/MPRL.pdf

Microplastics metadata template development

Arctic Monitoring & Assessment Programme

https://litterandmicroplastics.amap.no/

UN Environment Programme

https://www.unep.org/resources/report/microplastics

US Environmental Protection Agency

Water Quality Exchange Web Template Files

Note About Template Files

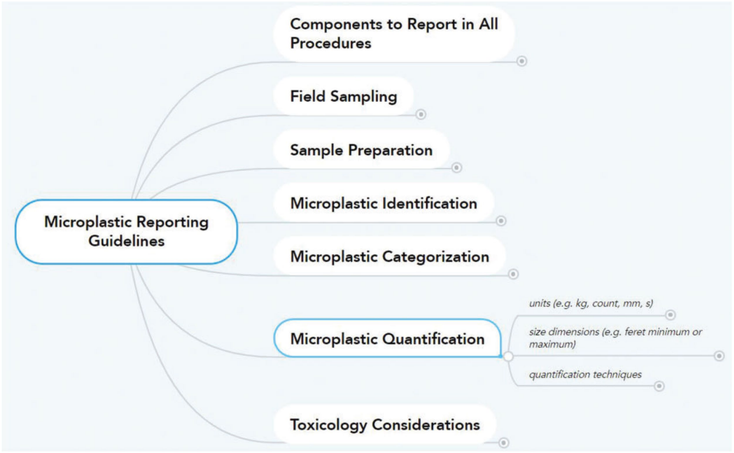
The templates are made up of two Microsoft Excel spreadsheet files that are meant to be used together to assist with data tracking and entry. The WQX Web Template Dictionary file provides guidance on the appropriate use of each data element, highlighting the data elements that are available in the WQX Web Template and showing the additional data elements that a user may add to the template. The dictionary covers data elements for all templates including:

- Physical-Chemical Results
- Biological Results
- Habitat Results, Activity Metrics and Indices
- Continuous Monitoring Results
- Lab Results

Many of the templates contain sample data that can be imported into WQX Web using the following import configurations (for more information regarding how to use Import Configuration files in WQX Web, please see the WQX web tutorials).

https://www.epa.gov/waterdata/water-quality-exchange-web-template-files

Academic Literature

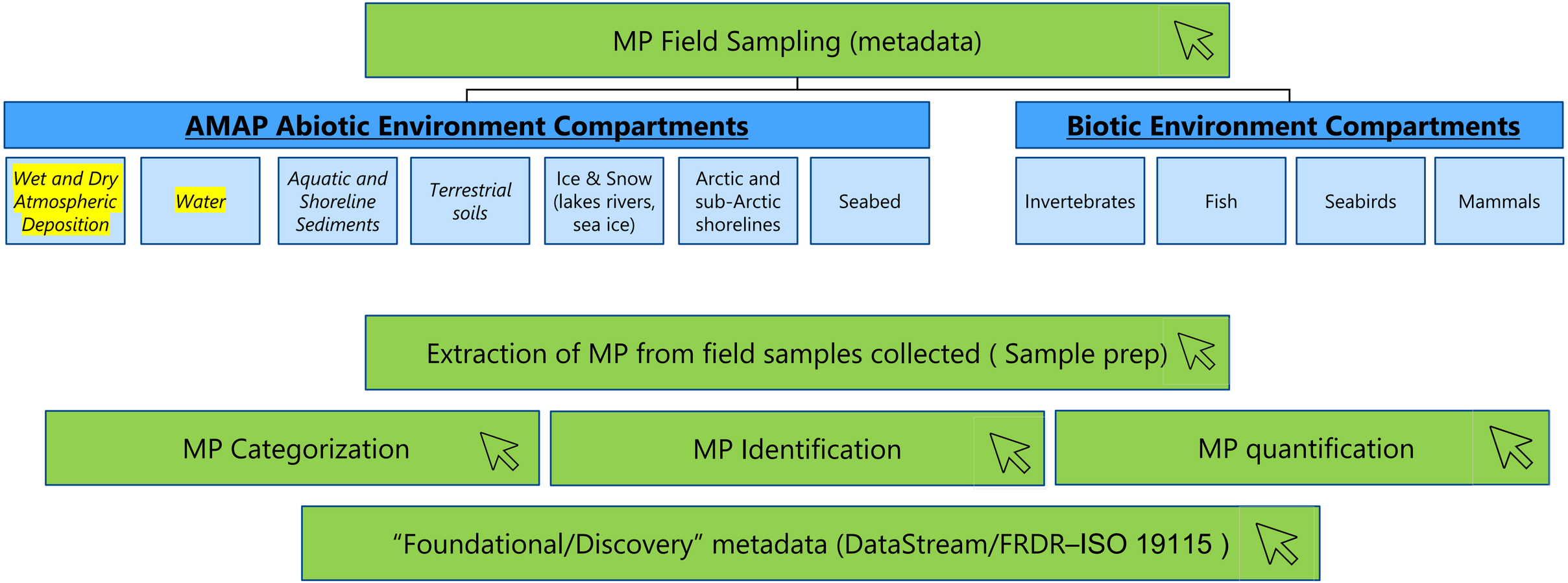


W. Cowger *et al. Appl. Spectr.* **2020**, *74*, 1066-1077.

	Number of elements	Percent of total
From DataStream & WQX	**36**	**21%**
From WQX	**33**	**19%**
New proposed	**101**	**59%**
Total elements from WQX	**69**	**41%**
Total elements in template	**170**	**100%**

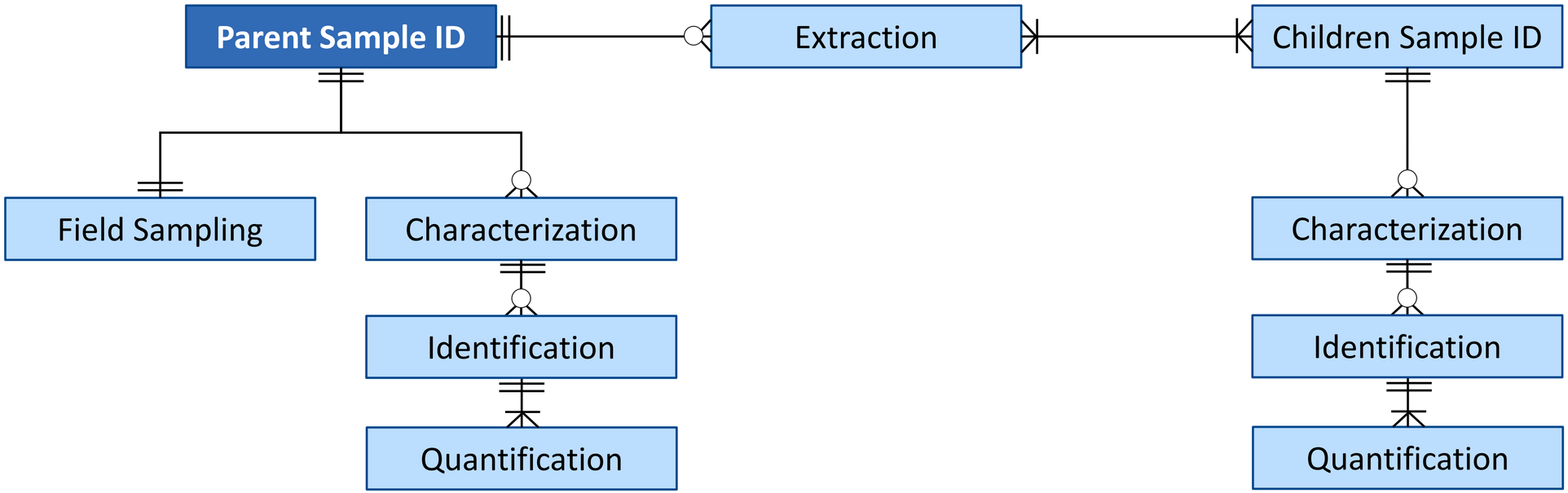
Putting it Together

"Use metadata are describing the actual content of a dataset and how it is encoded.
The purpose is to **enable the user to understand the data without any further communication**.
It describes content of variables using **standardized vocabularies**, units of variable, encoding of missing values, map projections etc"



Next Steps

- Review template (in excel) and provide your feedback on suite of variables identify
- Beta testing users



MP Field sampling (metadata)

- Monitoring Location
  - Location ID
  - Location name
  - Location latitude
  - Location longitude
  - Location horizontal Coordination reference system
    - Reach Length Measure — units
    - Net Type
    - Net Surface Area Measure — units
    - Boat Speed Measure — units
    - Current Speed Measure — units
  - Location type
- Activity Date/Time
  - Activity start date
  - Activity start time
  - Activity end date
  - Activity end time
- Media Name (a.k.a AMAP Compartment)
  - Volume of sample (Wet) — units
  - Mass of sample (Dry) — units
  - Activity type
- Activity Depth
  - Activity depth height measure
  - Activity depth height unit
- Sample Collection
  - Sample collection equipment
    - Device dimensions
      - Width
      - Length
      - Height
      - Aperture size
      - Mesh size
    - Deployment procedures
    - Filter type
  - Sample Collection method
    - If individual samples were pooled
    - Number of samples per site
- Sample Container Type
  - Material sample is transported in
  - solution sample is transported and stored in

Yellow text: mandatory on DataStream
White text: optional on DataStream
Red text: WQX only

Glassware cleaning method should be described

Report suppliers of labware, consumables, chemicals, and instruments

For labware, report relevant material properties (glass, plastic, etc.) and sizes/volumes

water and sediments | water samples | sediment samples | Cowger *et. al.*

MP Sample Preparation (metadata)

Sample Preparation Method

- Digestion
  - Duration time
  - Temperature of incubation
  - Composition of digestion solution
  - Concentration of digestion solution
  - Ratio of digestion fluid to sample
    - Units
- Density separation
  - Concentration of salt in mass per volume of the solution
  - Density of the solution
  - Type of salt used for fluid
  - Method used to remove salts from isolated microplastics
  - Duration of separation or settling
  - Device used for density separation
  - Sediment fluidization
  - Froth flotation
  - Separation funnel
  - Elutriation
- Filtration
  - Composition
  - Porosity
  - Diameter
- Oil separation
- Other developing new techniques

Conditional

Sample processing methods recommended for water

For labware, report relevant material properties (glass, plastic, etc.) and sizes/volumes

Where relevant, specify quantities of consumables used

Describe method used to clean glassware

Report method for filtering reagents

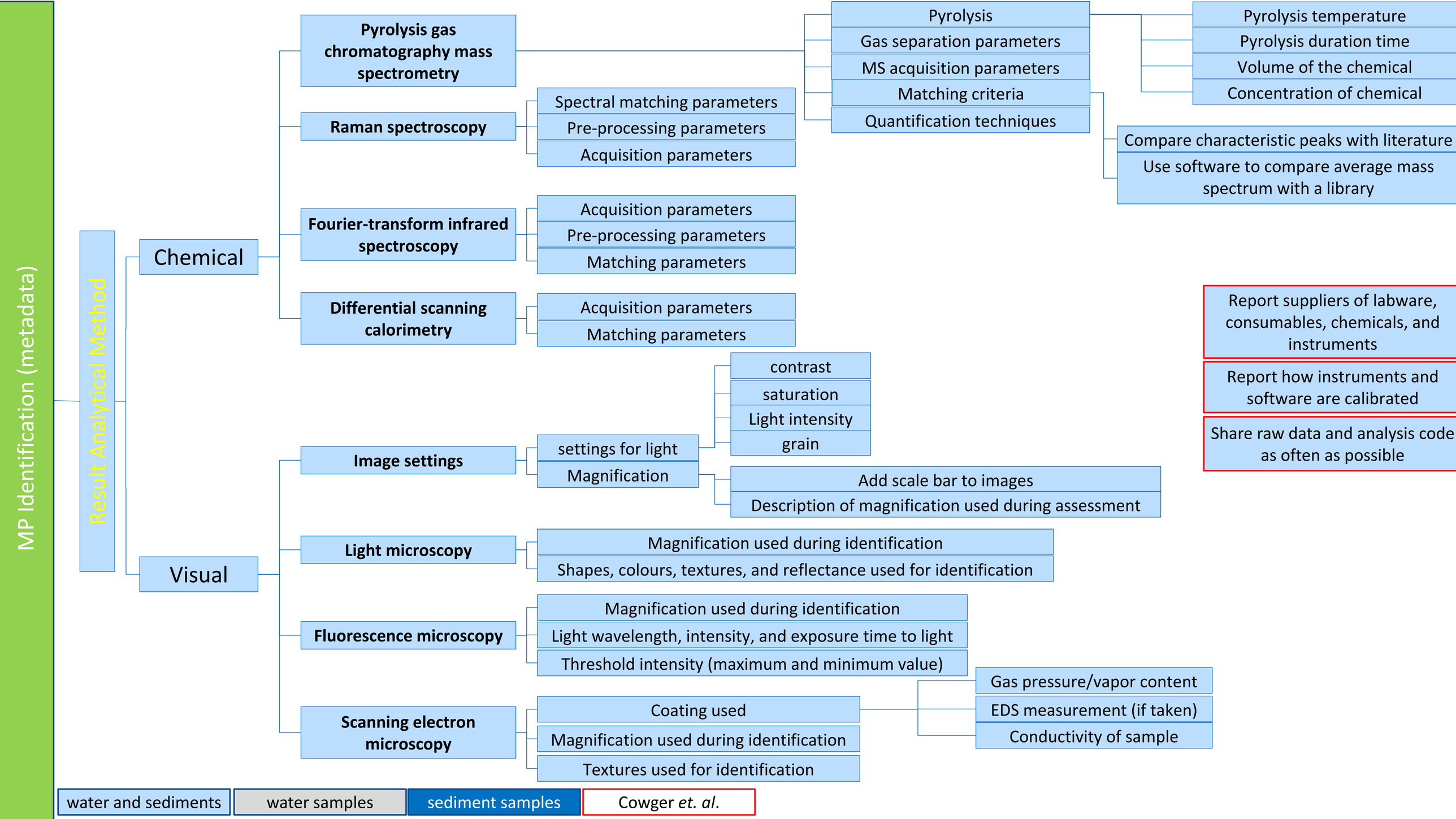
Material of clothing and gloves used in the lab should be described

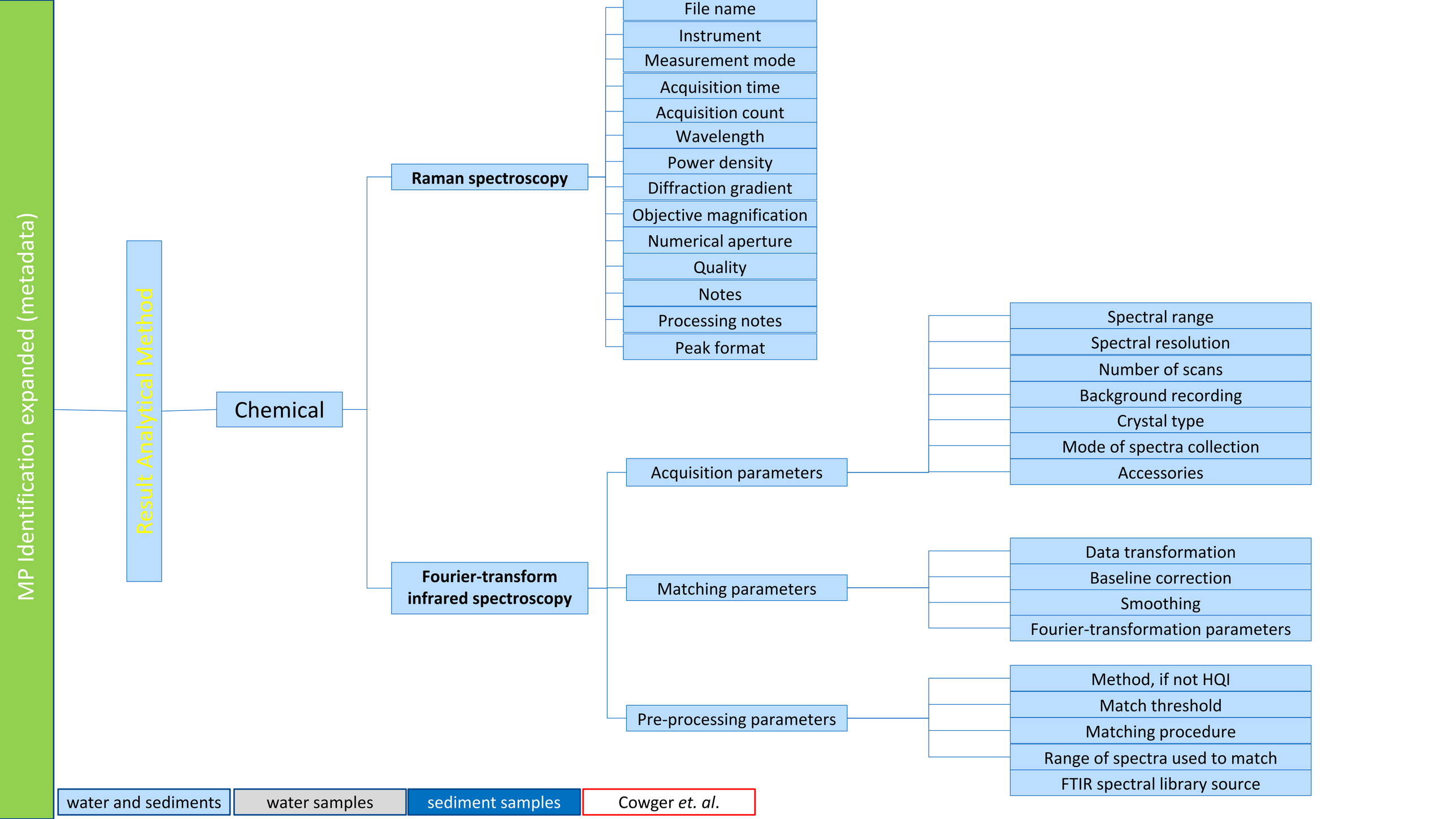
Report suppliers of labware, consumables, and chemicals.

Share raw data and analysis code as often as possible

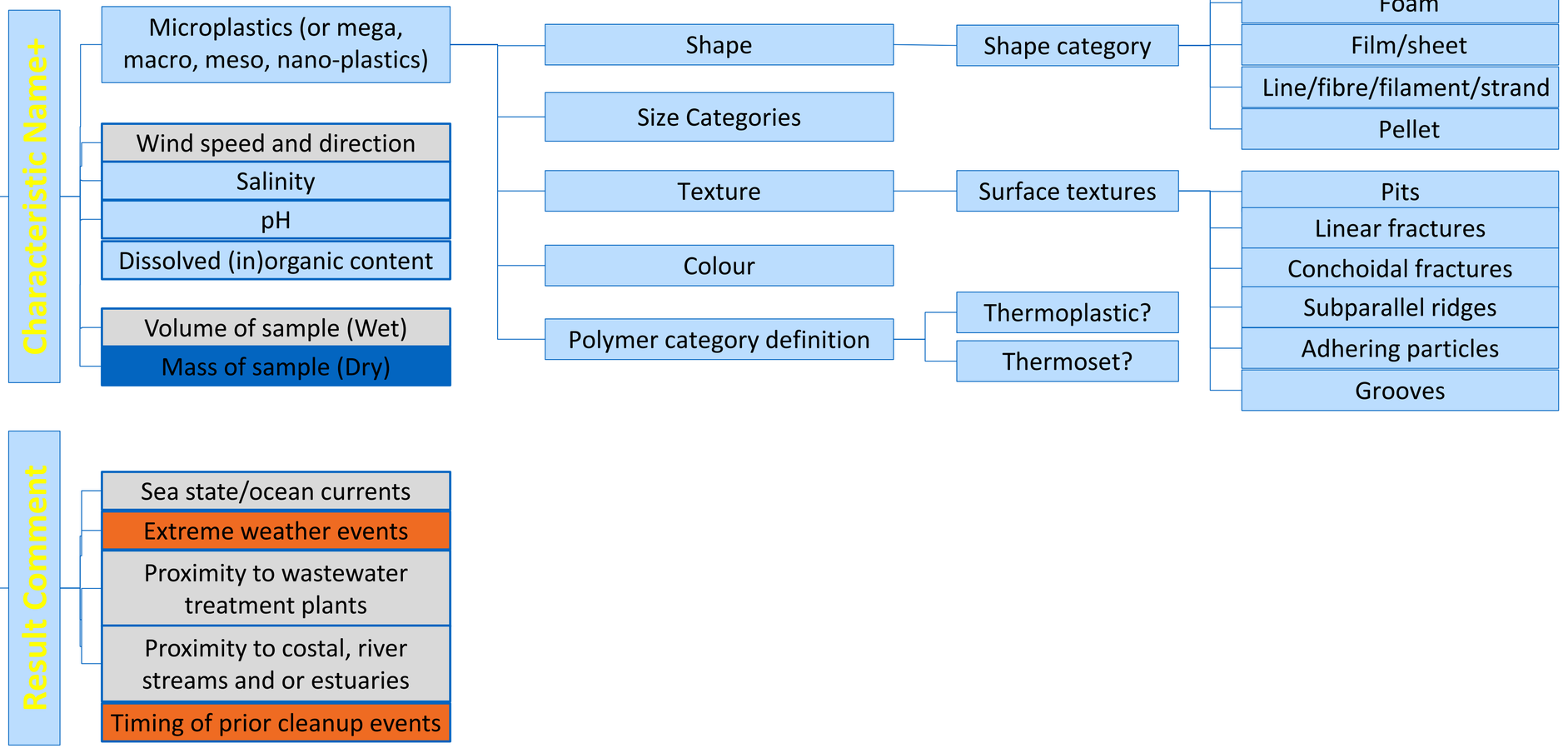
Describe contaminant devices used (e.g. flow hoods, glove bags)

water and sediments | water samples | sediment samples | Cowger *et. al.*


MP Identification (metadata)
Result Analytical Method
Chemical
Pyrolysis gas chromatography mass spectrometry
Pyrolysis
Gas separation parameters
MS acquisition parameters
Matching criteria
Quantification techniques
Pyrolysis temperature
Pyrolysis duration time
Volume of the chemical
Concentration of chemical
Compare characteristic peaks with literature
Use software to compare average mass spectrum with a library
Raman spectroscopy
Spectral matching parameters
Pre-processing parameters
Acquisition parameters
Fourier-transform infrared spectroscopy
Acquisition parameters
Pre-processing parameters
Matching parameters
Differential scanning calorimetry
Acquisition parameters
Matching parameters
Report suppliers of labware, consumables, chemicals, and instruments
Report how instruments and software are calibrated
Share raw data and analysis code as often as possible
Visual
Image settings
settings for light
contrast
saturation
Light intensity
grain
Magnification
Add scale bar to images
Description of magnification used during assessment
Light microscopy
Magnification used during identification
Shapes, colours, textures, and reflectance used for identification
Fluorescence microscopy
Magnification used during identification
Light wavelength, intensity, and exposure time to light
Threshold intensity (maximum and minimum value)
Scanning electron microscopy
Coating used
Gas pressure/vapor content
EDS measurement (if taken)
Conductivity of sample
Magnification used during identification
Textures used for identification
water and sediments
water samples
sediment samples
Cowger et. al.


MP Identification expanded (metadata)
Result Analytical Method
Chemical
Raman spectroscopy
File name
Instrument
Measurement mode
Acquisition time
Acquisition count
Wavelength
Power density
Diffraction gradient
Objective magnification
Numerical aperture
Quality
Notes
Processing notes
Peak format
Fourier-transform infrared spectroscopy
Acquisition parameters
Spectral range
Spectral resolution
Number of scans
Background recording
Crystal type
Mode of spectra collection
Accessories
Matching parameters
Data transformation
Baseline correction
Smoothing
Fourier-transformation parameters
Pre-processing parameters
Method, if not HQI
Match threshold
Matching procedure
Range of spectra used to match
FTIR spectral library source
water and sediments
water samples
sediment samples
Cowger et. al.

MP Categorization (metadata)



Standardized shape categories
[GESAMP: Pg. 11, Table 2.3]

Common sizes identified by measurable units
[GESAMP: Pg 10, Table 2.2]

Standardized texture descriptions
[Textures: Section 3]

12-colours classification (ISCC-NBS)
or
8-colour classification (EMODnet)
Method used to determine and classify colour should be discussed.
[GESAMP: Pg. 12]

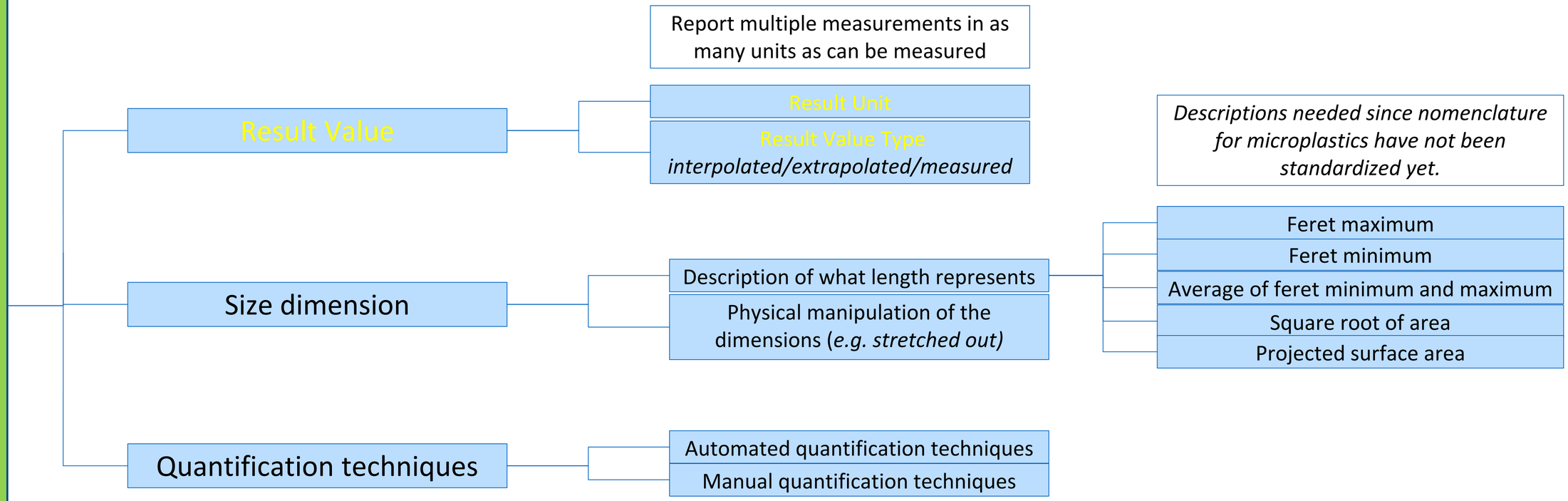
Standardized polymer definitions
[GESAMP: Pg. 6]

Is there a library of polymer choices that can used?

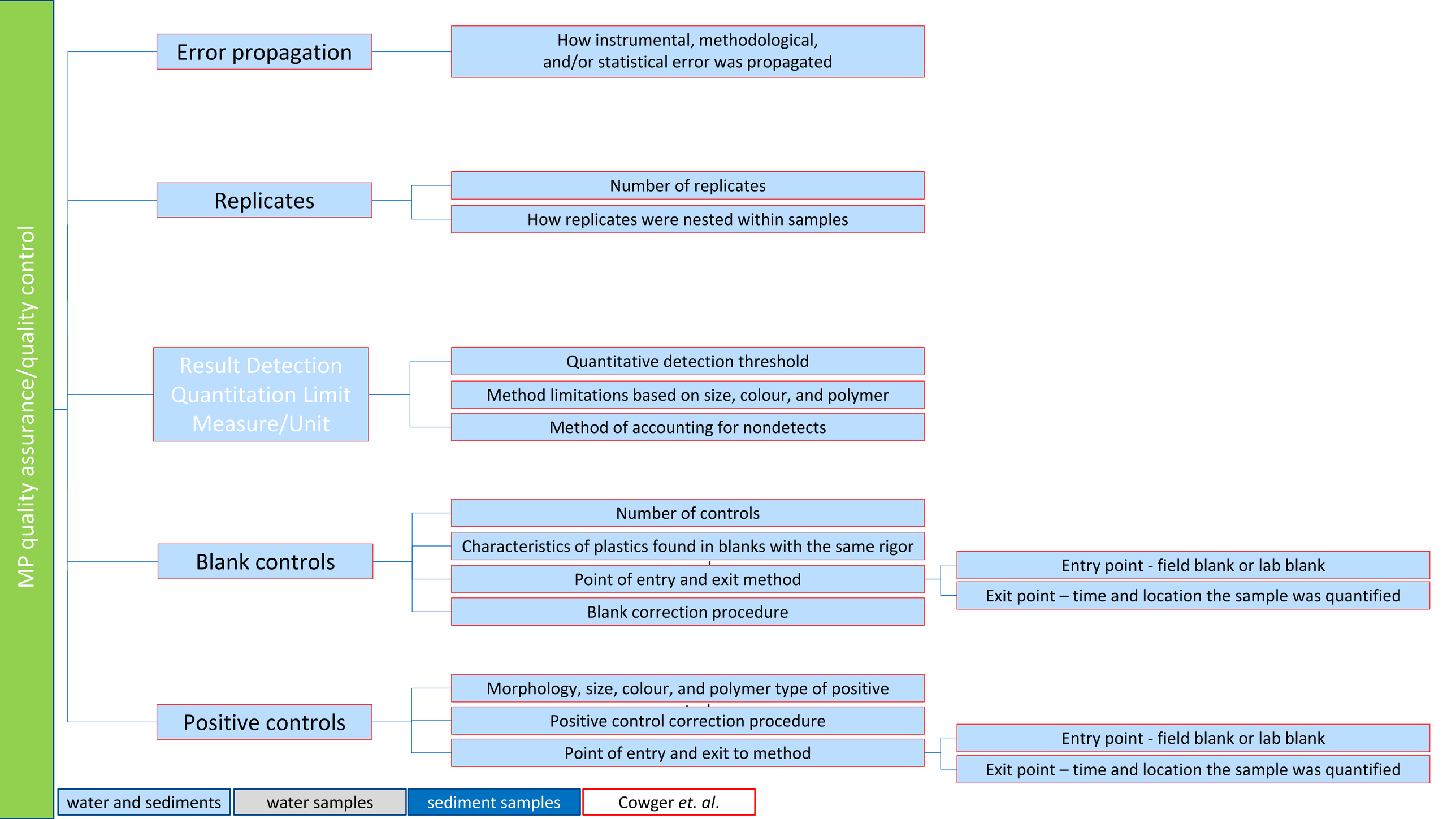
Textures: https://www.surfacesciencewestern.com/pdf/0909_mpb09_biesinger.pdf
GESAMP Guidance Document: http://www.gesamp.org/publications/guidelines-for-the-monitoring-and-assessment-of-plastic-litter-in-the-ocean

water and sediments | water samples | sediment samples | Cowger *et. al.*

MP quantification (metadata)



water and sediments | water samples | sediment samples | Cowger *et. al*.


MP quality assurance/quality control
Error propagation
How instrumental, methodological, and/or statistical error was propagated
Replicates
Number of replicates
How replicates were nested within samples
Result Detection Quantitation Limit Measure/Unit
Quantitative detection threshold
Method limitations based on size, colour, and polymer
Method of accounting for nondetects
Blank controls
Number of controls
Characteristics of plastics found in blanks with the same rigor
Point of entry and exit method
Entry point - field blank or lab blank
Exit point – time and location the sample was quantified
Blank correction procedure
Positive controls
Morphology, size, colour, and polymer type of positive
Positive control correction procedure
Point of entry and exit to method
Entry point - field blank or lab blank
Exit point – time and location the sample was quantified
water and sediments
water samples
sediment samples
Cowger *et. al*.

Discovery level (metadata)

Used to find relevant data once it is a repository

Suggest ISO 19115 metadata schema as DataStream and metadata fields also fits in well with:

- Federated Research Repository
- Polar Data Catalogue
- Open Government Metadata

Title of dataset	
Dataset Abstract	
Responsible person (s) & Affiliation & Orcid	
Study site name and coordinates	
Topic Category	
Keywords	at least five- must include “**microplastics**”
Data citation	
Time period data was collected	
Data security	
Funding	
Licensing & Attribution	
Maintenance Frequency	

Common to FRDR & DS

Data Stream only