Data Loading Checklist

Project:				
Data set:				
Completed by:				
Initiation date:				
Completion date:				

No.	Check	Y	N	N/A	Unk.	Details
1	Latitude and longitude are in a known coordinate system					
2	Target and actual coordinates are distinguished appropriately, and every sample has one or both					
3	Location identifiers have a 1:1 relationship to target coordinates					
4	Location name conflicts with location names for other studies have been resolved through assignment of study-specific location identifiers					
5	Sample identifiers are unique: no sample has multiple dates, depths, materials, or target locations					
6	When different values for the same attribute are recorded for the same sample (e.g., on different rows of the data set), consideration is given to both possibilities: a) only one set of attributes is correct, and b) the 'sample' is actually two or more samples. This consideration also applies to entities other than samples.					
7	Are ISM (incremental sampling methodology) samples from the same location and date clearly distinguished from sample splits (e.g., by different interpretive sample IDs)?					
8	Are there subsets of samples that should be grouped together, and distinguished from other subsets, using categories such as study element, sampling scheme, composite type, or collection SOP?					
9	All non-composite samples have a single collection date.					
10	All samples have an unambiguous material description.					
11	All water, soil, and sediment samples have sample depth information.					
12	All tissue samples have species information.					
13	Sampling information in the data source does not conflict with existing sample information that has undergone QA review.					

Data Loading Checklist

No.	Check	Y	N	N/A	Unk.	Details
14	Any existing locations or samples that are already in the database, possibly with different identifiers, have been identified and any discrepancies in identifiers or attributes have been resolved.					
15	Identifiers are consistently formatted. Spaces, hyphens, underscores, etc. are added or removed as necessary to ensure consistency. Original sample identifiers are retained in the "original_id" column of the database.					
16	When some data values are calculated from others, only the fundamental values are loaded. There may be exceptions; if so, they should be documented.					
17	Relationships among samples--collections, interpretive samples, analytical samples, containers, composites, and lab samples-- are all properly represented.					
18	The actual number of samples or observations is consistent with the planned number, if the latter is documented.					
19	Date/time values for different entities make logical sense. (For example, gear time in water should be before sample time.)					
20	Information that has systematically been recorded in comments has been loaded into appropriately specific columns in database tables.					
21	Water depths and sample depths are properly distinguished.					
22	Units are specified for all numeric values. If a new unit code must be added to IDB, the dimension and conversion factors for the new code have been checked by two independent reviewers who are fully conversant with how the conversion factors are used.					
23	Existing code lists have been thoroughly reviewed prior to adding new codes, and no duplicate codes have been added.					
24	All specifications in the Data Management Plan (DMP) and Data Manager's Manual (DMM) have been followed.					
25	All questions and issues relating to the data, and their resolutions, are documented in an issue log.					
26	Any required structural changes to the database have been reviewed by the database architect or delegate.					
27	The data loading operation has been scripted and documented, including references to issue log items as appropriate.					

Data Loading Checklist

No.	Check		Y	N	N/A	Unk.	Details
28	All new generally-applicable specifications and caveats arising from this loading process are documented in the DMM.						
29	Quality assurance checks of the loaded data have been completed in accordance with project requirements.						
30	Workflows and workflow steps have been recorded as required for the project and as specifed in the DMP.						