**S Table 1**. A subset of the DeepData database export file (5 fields of 49), showing how observations are distributed both across rows and columns, or a combination of both wide and long format in the data. Text in bold italic represent distinct observations, text in grey, repeated, redundant data. The ‘Analysis’ field represents column headings (from the data template) and is paired with the adjacent ‘Result’ field (therefore in ‘long’ format, unlike the rest of the data. The data were restructured to entirely wide format- ‘spread’ over separate columns, e.g. for ‘Ecology’ data in the ‘Category’ column, the ‘Nominal size Category’ (in ‘Analysis’) became a separate column (e.g. adjacent to the ‘Result’ column), with the entries as recorded in ‘Result’, i.e. ‘macro’

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **ContractorID** | **Category** | **TaxaID** | **Analysis** | **Result** |
| BGR1 | Ecology | 612 | Nominal size Category | **macro** |
| BGR1 | Ecology | 612 | Number of individuals | ***1*** |
| BGR1 | Organism Details | 612 | Sex | ***Unidentified*** |
| BGR1 | Taxonomist information | 612 | Taxonomist | ***Not Reported*** |
| BGR1 | Taxonomist information | 612 | Taxonomist E-mail | ***Not Reported*** |
| BGR1 | Taxonomist information | 612 | Taxonomist Institution | ***Not Reported*** |
| BGR1 | Taxonomy ID | 612 | Class | ***Polychaeta*** |
| BGR1 | Taxonomy ID | 612 | Identification Method | ***Not Reported*** |
| BGR1 | Taxonomy ID | 612 | Kingdom | ***Not Reported*** |
| BGR1 | Taxonomy ID | 612 | Morphotype | ***Not Reported*** |
| BGR1 | Taxonomy ID | 612 | Order | ***Not Reported*** |
| BGR1 | Taxonomy ID | 612 | Phylum | ***Annelida*** |
| BGR1 | Taxonomy information | 612 | Database Taxa ID | ***Not Reported*** |
| BGR1 | Taxonomy information | 612 | Taxonomic Database | ***Not Reported*** |
| BGR1 | Taxonomy information | 612 | Taxonomic Status | ***Uncertain*** |
| BGR1 | Ecology | 613 | Nominal size Category | **macro** |
| BGR1 | Ecology | 613 | Number of individuals | ***1*** |
| BGR1 | Organism Details | 613 | Sex | ***Unidentified*** |
| BGR1 | Taxonomist information | 613 | Taxonomist | ***Not Reported*** |
| BGR1 | Taxonomist information | 613 | Taxonomist E-mail | ***Not Reported*** |
| BGR1 | Taxonomist information | 613 | Taxonomist Institution | ***Not Reported*** |
| BGR1 | Taxonomy ID | 613 | Class | ***Polychaeta*** |
| BGR1 | Taxonomy ID | 613 | Identification Method | ***Not Reported*** |
| BGR1 | Taxonomy ID | 613 | Kingdom | ***Not Reported*** |
| BGR1 | Taxonomy ID | 613 | Morphotype | ***Not Reported*** |
| BGR1 | Taxonomy ID | 613 | Order | ***Not Reported*** |

.