**Interim Controlled Vocabulary Database**

**Goal**

Design, build and populate a simple controlled vocabulary database to develop a ‘seed’ list of controlled vocabularies that will ultimately be available: in *MCP metadata*; for use in faceted search lists of the AODN/IMOS Portal ; and by linked reference in delivered data services. Note that some controlled vocabularies will ultimately require their own data store (e.g., Feature Types should ultimately live in a Feature Catalogue, Geographic Names in a Gazetteer, People will need global UIDs and be part of a global registration system etc.). Eventually we need to use GeoNetwork and SISSVOC to manage and serve these Vocabs but for now we just need to get something up in-house that lets us start compiling the seed lists.

**Suggested Controlled Vocabulary Types**

* *Parameter Name* (roll our own)
* Collection Method
* *Platform* (roll our own)
* *Instrument* (roll our own)
* *Sampling Method* (roll our own)
* *Analysis Method* (roll our own)
* *Feature Type* (roll our own)
* Units Of Measure (choose a suitable, semantically-well defined existing source of UoM)
* Datum (roll our own)
* *Taxonomic Name* (agree to use a prioritised group of taxa name services/sources)
* *Organisation* (roll our own – but treated differently in Vocab Database – see below)
* People (Probably for IMOS purposes only – as there are now global UUIds for research staff which we should link into if we were going to do this in a serious fashion)
* *Geographic Location* (should be approached by partnering with GA using a national Gazetteer, however we could start gathering Geographic Locations for the purposes of pushing them to GA and perhaps the MarineRegions.org)
* *Country*

**Generic Table Structures For Controlled Vocabs**

**Vocabulary\_Register Table**

|  |  |  |  |
| --- | --- | --- | --- |
| **Column Name** | **Key** | **Description** | **Type** |
| Register\_Name | PK | Compact, human readable unique name for the register. | CharString |
| Register\_UID |  | A unique web addressable identifier for the Register (e.g.: http://vocab.aodn.org.au/def/parameter) | URL |
| Register\_Content\_Summary |  | General statement of the purpose of the register. | CharString |
| Register\_Language\_Name |  | Language the register is stored in (e.g. Welsh) | Char |
| Register\_Language\_Code |  | 3 Char code as specified in ISO 639-2. | Char |
| Register\_Language\_Country |  | 3 Char code as specified in ISO 3166-1 | Char |
| Register\_Language\_Encoding |  | Is an instance of MD\_CharacterSetCode (ISO 19115:2003, B.5.10. | Char |
| Register\_Version |  | This is the path to the versioned vocabulary file (e.g. http/vocab/aodn.org.au/def/platform/version1.0). This versioned URL is not used by default top reference current vocabularies. The URL is used instead for routinely accessing content from saved (stored) versioned copies of a vocabulary (in SKOS format) | (URI) URL |
| Register\_Date |  | Version date given to the ‘register version’ at the URI above. | Date |
| Date\_Of\_Last\_Change |  | Last date on which any item in the register had its status changed. | Date |

**Register\_Owner Table**

|  |  |  |  |
| --- | --- | --- | --- |
| **Column Name** | **Key** | **Description** | **Type** |
| Register\_Owner\_Id | PK | Id for the Register Owner | Integer |
| Person\_Id | FK | The person who is a register owner (e.g., Australian Ocean Data Network). | Integer |
| Organisation\_Name\_Id | FK | The organisation that the Person listed in Person\_Id is associated with (as a Vocabulary\_Owner) | Integer |
| Register\_Name | FK | Compact, human readable unique name for the register holding the vocabulary. | Charstring |

**Register\_Manager Table**

|  |  |  |  |
| --- | --- | --- | --- |
| **Column Name** | **Key** | **Description** | **Type** |
| Register\_Manager\_Id | PK | Id for the Register Manager | Integer |
| Person\_Id | FK | The person responsible for managing the register. | Integer |
| Organisation\_Name\_Id | FK | The organisation that the Person listed in Person\_Id is associated with (as a Vocabulary\_Manager) | Integer |
| Register\_Name | FK | Compact, human readable unique name for the register being managed. | Charstring |

**Term\_Proposer Table**

|  |  |  |  |
| --- | --- | --- | --- |
| **Column Name** | **Key** | **Description** | **Type** |
| Term\_Proposer\_Id | PK | Id for the Term Proposer | Integer |
| Person\_Id | FK | The person responsible for managing the register. | Integer |
| Organisation\_Name\_Id | FK | The organisation Id that the Person listed in Person\_Id is associated with (as a Term\_Proposer). | Integer |

**Vocabulary\_Term Table**

|  |  |  |  |
| --- | --- | --- | --- |
| **Column Name** | **Key** | **Description** | **Type** |
| Vocabulary\_Term\_Name | PK | A unique name given to the vocabulary Term. | CharString |
| Vocabulary\_Type\_Name | FK | The type of Vocabulary Term (e.g. Parameter, Instrument, Platform, Analysis Method) | CharString |
| Vocabulary\_Term\_Code |  | A unique code allocated to the Term (for use in an http:URI) | Integer |
| Vocabulary\_Term\_UID |  | A unique Web addressable identifier for this Term.  (e.g.: http://vocab.aodn.org.au/def/parameter/751) | Charstring |
| Vocabulary\_InstanceOf |  | Some Terms are instancesOf other Terms. In order to discern which Terms are instancesOf other Terms this flag should be set to ‘True’. This should then trigger parsing of the Internal\_Associated\_Terms Table in order to determine which Term, this particular Term is an instanceOf. | Boolean |
| Vocabulary\_Term\_Short\_Name |  | A short name given to the Term. |  |
| Vocabulary\_Term\_Ancilliary\_Information |  | Any other specific information that you want to record about the Term expressed as couplets of “ancilliary\_descriptor:value” separated by semi-colons if there is more than one descriptor (e.g., SerialNumber:22345). | CharString |
| Vocabulary\_Term\_Definition |  | A succinct definition for the named Term. | CharString |
| Alternative\_Term\_Expression |  | Alternative names given to the Term. Could just place a number of names in this field delineated by a comma (or we will need to create Alternate\_Term\_Expression table and associations) | CharString |
| Date\_Added |  | The date on which the Term was added to the register. | Date |
| Added\_By |  | The name of the eMII person who added the Term to the Register. | CharString |
| Term\_Proposer\_Id | FK | The Id and hence details of the person who proposed the Term. | CharString |
| Proposal\_Term\_Status |  | The proposal status of the Term (e.g., pending, tentative, final). | Char |
| Term\_Status |  | The status of the Term (e.g., notValid; valid; superseded; retired). |  |
| Date\_Status\_Amended |  | Date on which the Term Status was last changed. | Date |
| Reference\_Source \_Id | FK | The Unique Id for a record that describes the origins of the Vocabulary Term (if not invented by the AODN). If invented by the AODN this field would be null otherwise it has to have an entry. | Integer |
| Term\_Notes |  | Free text field in which a Term moderator can make notes about decisions taken regarding the Term. | CharString |
| Register\_Name | FK | Compact, human readable unique name for the register. | Charstring |
| Internal\_Relationship |  | Some Terms have relationships to other terms stored in the database. In order to discern which Terms have one or more of these relationships this flag should be set to ‘True’. This should then trigger parsing of the Internal\_Associated\_Terms Table. | Boolean |
| External\_Relationship |  | Some Terms have relationships to other terms stored not in the database but elsewhere on the Web. In order to discern which Terms have one or more of these relationships this flag should be set to ‘True’. This should then trigger parsing of the External\_Association\_Terms Table. | Boolean |

**Amendment Table (**A table which identifies proposed amendments to existing terms listed in the Vocabulary\_Term table).

|  |  |  |  |
| --- | --- | --- | --- |
| **Column Name** | **Key** | **Description** | **Type** |
| Amendment\_Id | PK | Id for the proposed amendment | Integer |
| Vocabulary\_Term\_Name | FK | A unique name given to the vocabulary Term to be amended. | Char |
| Vocabulary\_Term\_Code |  |  |  |
| Person\_Id | FK | The person responsible for suggesting the amendment. | Integer |
| Organisation\_Name\_Id | FK | The organisation Id that the Person listed in Person\_Id is associated with (as a Term Amender). | Integer |
| Register\_Name | FK | Compact, human readable unique name for the register in which the term belongs. | Charstring |
| Amendment\_Type |  | The type of amendment proposed (e.g. supersession or retirement) | Char |
| Amendment\_Request\_Date |  | Date on which amendment was requested. | Date |
| Amendment\_Approval\_Date |  | Date on which amendment was approved. | Date |

**Subject\_Term Table** (A table that re-casts a vocabulary term as a ‘Subject term’ for use in a relationship triplet of “Subject Term” – “Relationship Name” – “Object Term” for use in the Internal\_Associations\_Terms Table).

|  |  |  |  |
| --- | --- | --- | --- |
| **Column Name** | **Key** | **Description** | **Type** |
| Subject\_Term\_Id | PK | The unique Id for a vocabulary term cast as a subject term. | CharString |
| Vocabulary\_Term\_Name | FK | A unique name given to the vocabulary Term. | CharString |

**Object\_Term Table** (A table that re-casts a vocabulary term as an ‘Object term’ for use in a relationship triplet of “Subject Term” – “Relationship Name” – “Object Term” ” for use in the Internal\_Associations\_Terms Table).

|  |  |  |  |
| --- | --- | --- | --- |
| **Column Name** | **Key** | **Description** | **Type** |
| Object\_Term\_Id | PK | The unique Id for a vocabulary term cast as an object term. | CharString |
| Vocabulary\_Term\_Name | FK | A unique name given to the vocabulary Term. | CharString |

**Vocabulary\_Type Table** (Pick list of Vocabulary Types, i.e., Platform, Country, UoM, Instrument, SamplingMethod, AnalysisMethod and Feature Type. ‘People’ and ‘Organisation’ Type Terms are dealt with differently in the database – even though they are controlled vocabularies. ‘Taxonomy’ and ‘Geographic Location’ will be dealt with outside the scope of this database).

|  |  |  |  |
| --- | --- | --- | --- |
| **Column Name** | **Key** | **Description** | **Type** |
| Vocabulary\_Type\_Name | PK | The unique type name of a Vocabulary Term (e.g. Parameter, Instrument, Analysis Method). | CharString |
| Vocabulary\_Type\_Definition |  | A semantic and syntactic definition of the Vocabulary Type (i.e. an expression of the business rules involved in name construction). | CharString |

**Internal\_Association\_Terms Table** (Sometimes you want to be able to create an association between Terms stored within the database, for e.g., between a Parameter Name and a specific Unit of Measure or between Terms where there is a hierarchical relationship where one Term is considered to be an instance of another term . This can be achieved by creating named associations between Terms of interest using the Table below)

|  |  |  |  |
| --- | --- | --- | --- |
| **Column Name** | **Key** | **Description** | **Type** |
| Association\_Type\_Name | PK | The unique name of the Association Type (e.g.,isParameterUoM, isInstanceOf, isTypeOf, supercedes) | CharString |
| Subject \_Term \_Id | PK | Unique Id for a Vocabulary Term | Integer |
| Object\_Term \_Id | PK | Unique Id for a Vocabulary Term | Integer |

**External\_Association\_Terms Table** (Sometimes you want to be able to create an association, or mapping, between a Term in the database and another Term on the Web that is externally managed for e.g., between an internal Parameter Name and a Parameter Name from another external Web source).

|  |  |  |  |
| --- | --- | --- | --- |
| **Column Name** | **Key** | **Description** | **Type** |
| Association\_Type\_Name | PK | The unique name of the Association Type (e.g.,  skos:closematch; skos:exactmatch; skosnarrowmatch;  skos:broadmatch) | CharString |
| Vocabulary\_Term \_Name | PK | Unique Name for a Vocabulary Term in this database | Integer |
| External\_Term \_UID | PK | URL (UID) for a Vocabulary Term available externally on the Web. | CharString |

**Association\_Type Table**

|  |  |  |  |
| --- | --- | --- | --- |
| **Column Name** | **Key** | **Description** | **Type** |
| Association\_Type\_Name | PK | The unique name of the Association Type (e.g.,isParameterUoM, skos:exactmatch, isInstanceOf) | CharString |
| Association\_Type\_Name\_Description |  | A description of the Association Type. | CharString |

**Organisation\_Name Table**

|  |  |  |  |
| --- | --- | --- | --- |
| **Column Name** | **Key** | **Description** | **Type** |
| Organisation\_Name\_Id | PK | Unique Id for an Organisation | Integer |
| Organisation\_Name |  | The full Name of an Organisation participating in the AODN | CharString |
| Organisation\_Name\_Description |  | A definition of the Organisation | CharString |
| Organisation\_Acronym |  | An alternative current name for the Organisation | CharString |
| Organisation\_Address |  | Street Address of Organisation | CharString |
| Organisation\_Name\_Status |  | The status of an Organisation Name (e.g., currently\_valid, retired, superceeded). |  |
| Organisation\_Synonym |  | True or False. When an Organisation changes its name (and sometimes its function) over time, which happens frequently – particularly in the government, the Organisation\_Name\_Ids of related records should be stored in an Organisation\_Synonym Table. The various Name changes of the Organisation can then be tracked and used as synonyms. This, however, assumes a one to one correspondence between an old Organisation’s generic make-up and the new one). | Boolean |
| Organisation\_Notes |  | Free text field in which a moderator can make notes about decisions taken regarding the Organisation Name. | CharString |

**Organisation\_Synonym Table** (If Organisation\_Synonym in the Table above is set to “True”, related entries should be created in the Table below. Entries in this Table relate Organisations that have changed their names over time.)

|  |  |  |  |
| --- | --- | --- | --- |
| **Column Name** | **Key** | **Description** | **Type** |
| Organisation\_Synonym\_Id | PK | A Unique Id for a collection of related Organisation records that are referring essentially to the same Organisation through time. Creation of an Organisation Synonym record would be triggered if the status of an Organisation Name was “superceeded”. | Integer |
| Organisation\_Name\_Id | PK | Unique Id for an Organisation | Integer |
| Organisation\_Name |  | The full Name of an Organisation participating in the AODN | CharString |

**Person Table** (Not an ideal Table structure for storing people data – which is what Customer Relationship Management (CRM) systems are for. So this is a very simplified and condensed set of information)

|  |  |  |  |
| --- | --- | --- | --- |
| **Column Name** | **Key** | **Description** | **Type** |
| Person\_Id | PK | A Unique Id for a Person | Integer |
| Person\_Name |  | The official name the person is known by | CharString |
| Person\_Salutation |  | The salutation (e.g., Dr; Professor) | Char |
| Alternate\_Person\_Name\_Expression |  | Alternative names given to the Person. Could just place a number of names in this field delineated by a comma (or if this too slack we will need to create Alternate\_Name\_Expression table and associations) | CharString |
| Contact\_Email |  | Email address of Person | CharString |
| Contact Phone Number |  | Work Phone Number of Person | CharString |
| Contact Mobile Number |  | Mobile Phone Number of Person | CharString |

**Person\_Organisational\_Affiliation** (Organisations that the Person is affiliated with because it’s highly likely they will be associated with multiple Organisations)

|  |  |  |  |
| --- | --- | --- | --- |
| **Column Name** | **Key** | **Description** | **Type** |
| Person\_Id | PK | A Unique Id for a Person | Integer |
| Organisation\_Name\_Id | PK | A Unique Id for an Organisation | Integer |
| Affiliation\_Type\_Name | PK | A Unique Id for an Affiliation Type | Charstring |

**Affiliation Type** (The various Types of Affiliation a Person has with an Organisation)

|  |  |  |  |
| --- | --- | --- | --- |
| **Column Name** | **Key** | **Description** | **Type** |
| Affiliation\_Type\_Name | PK | Unique name for an Affiliation Type (e.g., employedBy; manages; memberOf) | Charstring |
| Affiliation\_Description |  | A description of the Affiliation Type | CharString |

**Reference\_Source Table** (The fields in this Table could sit in the main Vocabulary Term Table if desired)

|  |  |  |  |
| --- | --- | --- | --- |
| **Column Name** | **Key** | **Description** | **Type** |
| Reference\_Id | PK | Unique Id for a Reference Record | Integer |
| Citation\_String |  | Reference Style citation crediting the source for the Vocabulary Term | CharString |
| Online\_Reference\_Resource |  | A URL to the cited reference | URL |
| Organisation\_Name\_ID |  | A Unique Id for an Organisation | Integer |

**Classification\_Scheme\_Association Table** (This Table associates a Classification Scheme and a Category in that Scheme. Category terms can be nested hierarchically and be used in more than one scheme).

|  |  |  |  |
| --- | --- | --- | --- |
| **Column Name** | **Key** | **Description** | **Type** |
| Classification\_Scheme\_Id | PK | Unique Id for a Classification Scheme | Integer |
| Classification\_Scheme\_Category\_Id | PK | Unique Id for a Classification Scheme Category | Integer |
| Classification\_Scheme\_Category\_Label |  | A short human readable name identifying the category | Charstring |
| Hierarchy\_Level |  | Indicates level of nesting of category. If level =1 category is top category, level = 2 or greater indicates a nested category. | Integer |
| Parent\_Category\_Code | PK | This is the code for the category that is the parent (immediate category above the category being described). If level = 1 this field is null. | Integer |
| Parent\_Category\_Name |  | This is the name for the category that is the parent (immediate category above the category being described). If level = 1 this field is null. | URL |
| Parent\_Category\_Label |  | This is the short human readable name for the category that is the parent (immediate category above the category being described). If level = 1 this field is null. | Charstring |

**Classification\_Scheme\_Category Table**

|  |  |  |  |
| --- | --- | --- | --- |
| **Column Name** | **Key** | **Description** | **Type** |
| Classification\_Scheme\_Category\_Id | PK | Unique Id for a Classification Scheme Category | Integer |
| Classification\_Scheme\_Category\_Name |  | The URL of a Category used in a Classification Scheme | URL |
| Classification\_Scheme\_Category\_Label |  | A short human readable name identifying the category | CharString |
| Classification\_Scheme\_Category\_Description |  | A description of the Category used in a Classification Scheme | CharString |
| Date\_Added |  | The date the category was added to the system for publishing | Date |
| Classification\_Scheme\_Id | FK | Unique Id for a Classification Scheme | Integer |

**Classification\_Scheme Table**

|  |  |  |  |
| --- | --- | --- | --- |
| **Column Name** | **Key** | **Description** | **Type** |
| Classification\_Scheme\_Id | PK | Unique Id for a Classification Scheme | Integer |
| Classification\_Scheme \_Name |  | The URL of a Classification Scheme | URL |
| Classification\_Scheme\_Title |  | Short human readable title for the scheme | CharString |
| Classification\_Scheme \_Description |  | A description of the Classification Scheme | CharString |
| Person\_Id |  | The person in the organisation who is responsible for publishing the scheme | Integer |
| Organisation\_name\_Id |  | The organisation which is responsible for creating the scheme | Integer |
| Data\_Added |  | The date on which the scheme was added | Date |

**Term\_Category\_Classification Table** (This Table associates a Classification Scheme, a Category in that Scheme and a Vocabulary Term so that Vocabulary Terms can be grouped, or classified according to different categories and schemes. A vocab term can be associated with more than one category within a single scheme).

|  |  |  |  |
| --- | --- | --- | --- |
| **Column Name** | **Key** | **Description** | **Type** |
| Vocabulary\_Term\_Code | PK | Unique Id for a Vocabulary Term that is associated with a given Classification Scheme and a specific Category in that Scheme | Integer |
| Vocabulary\_Term\_UID |  | A unique Web addressable identifier for this Term.  (e.g.: http://vocab.aodn.org.au/def/parameter/751) | Charstring |
| Vocabulary\_Term\_Name |  | Unique Name for a Vocabulary Term that is associated with a given Classification Scheme and a specific Category in that Scheme | Integer |
| Classification\_Scheme\_Category\_Id | PK | Unique Id for a Classification Scheme Category | Integer |
| Classification\_Scheme\_Id | PK | Unique Id for a Classification Scheme | Integer |
| Date\_Added\_To\_association |  | The date on which the Term was added to the Scheme | Date |