ATLAS OF LIVING AUSTRALIA

Micro-organism Data Sharing Format

Author(s):	Branford, Matt (CES, Black Mountain)
Version:	v2.1
Creation Date:	12 December 2011

Table of Contents

Document Control Information

Document Location:	Available from http://www.ala.org.au/about/program-of-projects/biolomics/
UNC Path:	\\nexus\ento\Themes\1173-Biodiversity\Current Projects\ALA - Atlas of Living Australia\ALA Project Management 09 onwards\07. Project Work Packages & Deliverables\07.05. Collections Management\BioloMICS\AMRiN Shared Data Layout v2.1.docx

Document Change History:

Version	Date of Revision	Author	Change Description
V2.1	12/12/2011	Matt Branford	Change mandatory elements – 'kingdom' out, 'basisOfRecord' in Add chapter 2.8 'Recommended Minimum Data' Changed several field labels to use Darwin Core name where one exists (to better support 'single file' minimum data scenario). Previous field names are noted in the text. Change title from 'AMRIN Data Sharing Format' to 'Micro-organism Data Sharing Format'

Version	Date of Revision	Author	Change Description
v 2.0	22/9/2011	Matt Branford	Merge ‘imageURL’ and ‘imageFileName’ to ‘identifier’ Add ‘coordinatePrecision’ Add ‘eml.xml’ to zip file diagram Change all DWC URLs to reflect those actually used in resolution, not the ‘guide to use’
v 1.00	5/4/2011	Matt Branford	Approved version (nil content changes)
v 0.05	10/3/2011	Matt Branford	Updates from internal review
v 0.04	8/3/2011	Matt Branford	Updates from CHACM Technical Group review
v 0.03	19/1/2011	Matt Branford	Minor amendments from internal review
v 0.02	18/1/2011	Matt Branford	Use regular template, complete DWC archive description
v 0.01	17/1/2011	Matt Branford	Initial Draft

1 Introduction

1.1 Purpose

This document describes the format for sharing data with the Australian Microbial Resources Information Network (AMRiN) and Atlas of Living Australia.

1.2 Related Documents

The following table lists documents related to this document and that are to be used in conjunction with this document.

AMRiN Requirement Statement – describes the scope of data and search available under specially authorised and public access, and basis of the layout definition. Available from http://www.ala.org.au/about/program-of-projects/biolomics/
ALA Micro-organism Implementation Plan - Project plan under which this document is being developed Available from http://www.ala.org.au/about/program-of-projects/biolomics/
BioloMICS Implementation Proposal - ALA proposal to enhance Australian microorganism community's data management, and redevelop the AMRiN website. As approved by CHACM. Available from http://www.ala.org.au/about/program-of-projects/biolomics/

1.3 Layout Rationale and Design Decisions

This format balances competing factors.

1.3.1 Use of Microbial Common Language (MCL)

CHACM determined that ALA should consider use of MCL. Several AMRIN requirements involve data elements that are not represented in MCL. The format:

- References MCL property names and definitions wherever possible
- References MCL association/relationship names and definitions wherever possible.

The net effect is that many elements and concepts are semantically identical to corresponding MCL components.

1.3.2 Support for sharing different amounts of data

Data made available for sharing by a collection at a point in time depends on:

- Local policy decisions to share or not share particular elements
- Whether the information is held electronically when the data is shared.

The format:

- Has a minimal set of mandatory elements
- Includes optional elements that allow for all items in the AMRiN Data Requirement Statement

- Allows collections to provide minimum/mandatory data initially, and include optional elements later on.

Please note that AMRiN/ALA can also accept any additional Darwin Core fields that are not listed in this document. The full set of Darwin Core terms are available at http://rs.tdwg.org/dwc/terms/index.htm.

1.3.3 Re-use existing ALA components

ALA has already developed formats for floral and faunal collection data, for use in sharing 'public' and 'restricted' data. The same approach and components are to be used wherever possible.

The net impact is:

- the format is based on a "Darwin Core Archive extension" approach
- improved reliability – re-used components are already developed and proven
- significant benefit to those institutions using the same system for all their collection types (same format style minimises development/maintenance effort).

1.3.4 Facilitate adoption where IT Funding/Skills not available

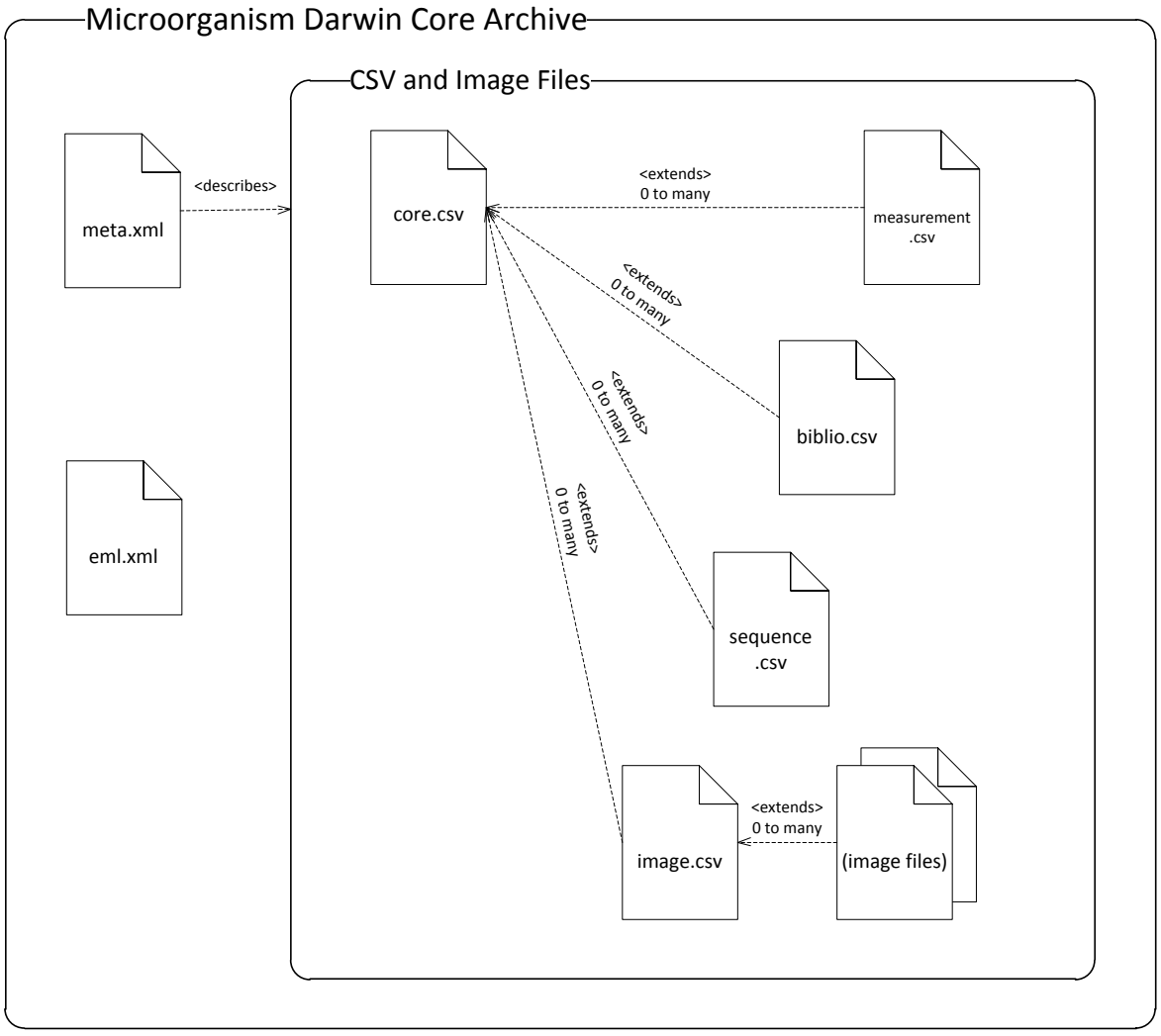
During requirement gathering, many collections indicated they would not be able to implement complicated formats. This is another driver for using the CSV-based 'Darwin Core Archive' approach.

2 Darwin Core Archive Definition

2.1 Overview

ALA uses the same format style for sharing microorganism data as for floral and faunal collections. The format is a Darwin Core Archive extension. http://rs.tdwg.org/dwc/terms/guides/text/index.htm#implement has more information about Darwin Core extensions.

In essence, the 'archive' is a zip file containing specially formatted CSV files and associated images and contig files, and an xml file that describes the structure.



The layouts of each file are detailed in the rest of this document. An overview of each file is:

- eml.xml – describes the source of the data
- meta.xml – static (unchanging) file that describes the zip file contents
- core.csv – Minimum/Mandatory and optional fields covering:

  - what the microorganism is
  - terms by which it is known or referenced by
  - which collection the specimen is held in
  - information about the sample
  - information about the medium
  - oxygen and temperature relationship
  - catalog and availability information
- characteristics.csv – optional fields about observations or characteristics other than oxygen and temperature relationship
- biblio.csv – optional fields about publications or papers or other material related to the microorganism
- sequences.csv – optional fields about DNA or Protein sequences of the specimen
- images.csv – optional fields about images of the specimen

2.2 Creating an Archive

To create the relevant Darwin Core Archive, collections should:

- Determine which elements they will share
- Create CSV reports that conform to relevant file layout definitions
- Zip together the CSV reports, any shared sequence files, any shared images and the meta.xml file
- Upload the zip file to ALA (through http://www.ala.org.au/share/share-data/ or similar).

ALA staff can assist collections to automate processes at every step.

2.3 Core File Layout

Mandatory fields are labelled and highlighted in red.

Field name	Description	Data Rules
catalogNumber *(previously strainNumber)*	The strain number. The institution or collection's unique internal reference number for the specimen. Used by AMRiN and ALA to link records within the "Darwin Core Archive" file format. For example "CBS 14". MCL term: http://www.straininfo.net/ns/mcl/2.0/strainNumber DWC term: http://rs.tdwg.org/dwc/terms/catalogNumber	Mandatory free text Must be unique within collectionCode
kingdom	The type of microorganism. For example, "Plantae", "Fungi" or "Chromista". Values should be as per classification in a recognised register such as ZooBank, PlantBank, MycoBank or ICTVdb – any register operating under the auspices of: • ICZN (International Committee of Zoological Nomenclature) • ICBN (International Committee of Botanical Nomenclature • ICSP (International Committee on Systematics of Prokaryotes) • ICTV (International Committee on Taxonomy of Viruses). DWC term: http://rs.tdwg.org/dwc/terms/kingdom	Free text
scientificName *(previously speciesName)*	Describes the species name. Species name contains genus and species epithet, and if applicable, the subspecies epithet. MCL term: http://www.straininfo.net/ns/mcl/2.0/speciesName DWC term: http://rs.tdwg.org/dwc/terms/scientificName	Mandatory free text
institutionCode	The name (or acronym) in use by the institution having custody of the object(s) or information referred to in the record. The value should also be represented in the AMRiN/ALA "Collectory" list. For example, "National Pathology Institute", "FMNH", "AKN-CLO".	Mandatory free text

	DWC term: http://rs.tdwg.org/dwc/terms/institutionCode	
collectionCode	The name, acronym, code, or initials identifying the collection or data set from which the record was derived. The value should also be represented in the AMRiN/ALA "Collectory" list. If not supplied, ALA/AMRiN will default the institutionCode value. For example, "Bacteria", "ANBC", "42". DWC term: http://rs.tdwg.org/dwc/terms/collectionCode	Free text
basisOfRecord	The specific nature of the data record. For example, "LivingSpecimen". DWC term: http://rs.tdwg.org/dwc/terms/basisOfRecord	Mandatory Values:"LivingSpecimen", "Taxon", "Occurrence", "PreservedSpecimen", "FossilSpecimen", "HumanObservation", "MachineObservation", null.
otherCatalogNumbers *(previously otherStrainNumbers)*	A list of other, equivalent strain numbers, separated by semicolons. For example, "ATCC 128;IFO 278;PYCC 937;SCHLEIN". MCL term: http://www.straininfo.net/ns/mcl/2.0/otherStrainNumbers DWC term: http://rs.tdwg.org/dwc/terms/otherCatalogNumbers	Free text
typeStatus *(previously typeDescription)*	Status of the strain/specimen (Information of the taxonomic status of the strain/specimen. Is a type, neo-type, etc and if so, of which taxa. For example, "Neotype of Rhodotorula rubra (Demme) Lodder (designated Lodder & Kreger-van Rij 1952)" MCL term: similar to http://www.straininfo.net/ns/mcl/2.0/typeStrainOf DWC term: http://rs.tdwg.org/dwc/terms/typeStatus	Free text
history	History of the strain/specimen, showing culture and deposit events. For example, "CBS -> IFO -> NITE", which reads as the strain was isolated at CBS, deposited to IFO, then deposited to NITE. MCL term: http://www.straininfo.net/ns/mcl/2.0/history	Free text

occurrenceRemarks *(previously comments)*	Remarks or comments about the strain or culture. For example, “Subject of Australian press reports December 2010”. MCL term: http://www.staininfo.net/ns/mcl/2.0/comments DWC term: http://rs.tdwg.org/dwc/terms/occurrenceRemarks	Free text
sampleOccurrenceType	Origin of the sample. Location information for Clinical records should not be provided for display on maps for public users. Values are: • Clinical - the sample was taken from a human patient or subject • Environmental - the sample was taken from a biome • Veterinary - the sample was taken from livestock, pets or wild animals.	Values: “Clinical”, “Veterinary”, “Environmental”, null
habitat *(previously sampleTextDescription)*	Textual description of habitat where sample was taken. Details of the isolation substrate. For example, “Isolated from fruit of Coumarouna punctata (tonka-bean)”. MCL term: http://www.staininfo.net/ns/mcl/2.0/sampleHabitat DWC term: http://rs.tdwg.org/dwc/terms/habitat	Free text
sampleSource	The source or host the sample was taken from. For example, “chicken”, “horse”, “human”, “mud”.	Free text Controlled vocabulary - new terms and synonyms will be flagged when uploading data
sampleSpecificSite	The specific site the sample was taken from. For example, “fruit”, “egg”, “dairy product”, “meat”, “blood”, “urine”, “mucus”. MCL term: similar to http://www.staininfo.net/ns/mcl/2.0/sampleHabitatEnvoTerm	Free text Controlled vocabulary - new terms and synonyms will be flagged when uploading data
country *(previously sampleLocationCountry)*	Country where sample was taken. For example, “Australia”. MCL term: http://www.staininfo.net/ns/mcl/2.0/sampleLocationCountry	Value must be listed at ISO3166 (including ISO3166-3, formerly used country names).

	DWC term: http://rs.tdwg.org/dwc/terms/country	
stateProvince	The name of the next smaller administrative region than country (state, province, canton, department, region, etc.) in which the Location occurs. For example, “ACT”, “NSW”, “NT, “QLD”, “SA”, “TAS”, “VIC”, ‘WA”. Value should be listed at ISO3166-2 (Country subdivision code). DWC term: http://rs.tdwg.org/dwc/terms/stateProvince	Free text Controlled vocabulary - new terms and synonyms will be flagged when uploading data
locality *(previously sampleLocationPlace)*	Locality of origin/where collected. Do not include patient addresses. For example, “Little Para River, Harry Bowey Reserve”. MCL term: http://www.straininfo.net/ns/mcl/2.0/sampleLocationPlace DWC term: http://rs.tdwg.org/dwc/terms/locality	Free text
geodeticDatum	The ellipsoid, geodetic datum, or spatial reference system (SRS) upon which the geographic coordinates given in decimalLatitude and decimalLongitude as based. For example, "EPSG:4326", "WGS84", "NAD27", "Campo Inchauspe", "European 1950". DWC term: http://rs.tdwg.org/dwc/terms/geodeticDatum	Free text
verbatimLatitude *(previously sampleVerbatimLatitude)*	Latitude at which sample was taken. MCL term: http://www.straininfo.net/ns/mcl/2.0/sampleLat DWC term: http://rs.tdwg.org/dwc/terms/verbatimLatitude	Free text
verbatimLongitude *(previously sampleVerbatimLongitude)*	Longitude at which sample was taken. MCL term: http://www.straininfo.net/ns/mcl/2.0/sampleLong DWC term: http://rs.tdwg.org/dwc/terms/verbatimLongitude	Free text
decimalLatitude *(previously sampleDecimalLatitude)*	Latitude at which sample was taken. The geographic latitude (in decimal degrees, using the spatial reference system given in geodeticDatum) of the geographic center of a Location. Positive values are north of the Equator, negative values are south of it. Legal values lie between -90 and 90, inclusive. DWC term: http://rs.tdwg.org/dwc/terms/decimalLatitude	Number

decimalLongitude *(previously sampleDecimalLongitude)*	Longitude at which sample was taken. The geographic longitude (in decimal degrees, using the spatial reference system given in geodeticDatum) of the geographic center of a Location. Positive values are east of the Greenwich Meridian, negative values are west of it. Legal values lie between -180 and 180, inclusive. DWC term: http://rs.tdwg.org/dwc/terms/decimalLongitude	Number
coordinatePrecision	A decimal representation of the precision of the coordinates given in the decimalLatitude and decimalLongitude. DWC term: http://rs.tdwg.org/dwc/terms/coordinatePrecision	Number
sampleAlt	Altitude at which sample was taken, including units. Depth is represented as a negative altitude. For example: • "1200m" • "-5m" MCL term: http://www.straininfo.net/ns/mcl/2.0/sampleAlt DWC term: similar to http://rs.tdwg.org/dwc/terms/verbatimElevation and http://rs.tdwg.org/dwc/terms/verbatimDepth	Free text
recordedBy *(previously sampleCollector)*	Name of sample collector MCL term: http://www.straininfo.net/ns/mcl/2.0/sampleCollector DWC term: http://rs.tdwg.org/dwc/terms/recordedBy	Free text
eventDate *(previously sampleDate)*	Date when sample was taken. Preferred format is YYYY-MM-DDThh:mm:ss±hhmm For example: • "2010-04-01T12:00:00+1000" – sample was taken April 1 2010 at noon AEST • "2010-04-01" – sample was taken April 1 2010. MCL term: http://www.straininfo.net/ns/mcl/2.0/sampleDate DWC term: http://rs.tdwg.org/dwc/terms/eventDate	Datetime ISO 8601

identifiedBy	Name of the person who identified the strain. MLC term: similar to http://www.straininfo.net/ns/mcl/2.0/isolator DWC term: http://rs.tdwg.org/dwc/terms/identifiedBy	Free text
biohazardLevel	The CDC/EU Biosafety/Pathogen/Protection Level for the microorganism. For example, “1” (causes only mild disease to humans, or is difficult to contract via aerosol in a lab setting).	Values: “1”, “2”, “3”, “4”, null
mediumName	Common name of culture medium. Example: 'Nutrient Agar (Oxoid CM3)' MCL term: http://www.straininfo.net/ns/mcl/2.0/mediumName	Free text
mediumDescription	Full description (including list of ingredients) of culture medium preparation. MCL term: http://www.straininfo.net/ns/mcl/2.0/mediumDescription	Free text
dateLastChecked	Date and time the culture was last checked and confirmed viable. Preferred format is YYYY-MM-DDThh:mm:ss±hhmm For example: • “2010-04-01T12:00:00+1000” – culture was checked April 1 2010 at noon AEST • “2010-04-01” – culture was checked April 1 2010.	Datetime ISO 8601
actualGrowthTemperature	The temperature the specimen was grown at (degree Celsius, do not include unit). Similar to MCL term: http://www.straininfo.net/ns/mcl/2.0/optimalGrowthTemperature	Number
minimalGrowthTemperature	Minimal temperature (degree Celsius, do not include unit) necessary to observe growth on culture medium. MCL term: http://www.straininfo.net/ns/mcl/2.0/minimalGrowthTemperature	Number
optimalGrowthTemperature	Temperature (degree Celsius, do not include unit) at which optimal growth on culture medium can be observed. MCL term: http://www.straininfo.net/ns/mcl/2.0/optimalGrowthTemperature	Number

maximalGrowthTemperature	Maximal temperature (degree Celsius, do not include unit) at which growth on culture medium can be observed. MCL term: http://www.straininfo.net/ns/mcl/2.0/maximalGrowthTemperature	Number
oxygenRelationship	One of AE, MA, FAN, AT, MAT, AN: • AE aerobic (100% air) • MA microaerophilic (5% air) • FAN facultative aerobic • AT aerotolerant (prefers anaerobic conditions for good growth, but tolerates aerobic condtions) • MAT microaerotolerant (anaerobic bacteria tolerating microaerophilic conditions) • AN anaerobic See Bacterial MINE, OXR. MCL term: http://www.straininfo.net/ns/mcl/2.0/oxygenRelationship	Values: “AE”, “MA”, “FAN”, “AT”, “MAT”, “AN”, null
catalogURL	Intended to support requests or purchases of a specimen. Link to online order form for the strain. MCL term: http://www.straininfo.net/ns/mcl/2.0/catalogURL	URL
availabilityRemarks	Intended to support requests or purchases of a specimen. Information about how to order the strain if there is no online order form. For example, “Available on request, email dr.hofstedter@myco.edu.au for details.”	Free text

2.4 Sequences File Layout

Field name	Description	Data Rules
catalogNumber *(previously strainNumber)*	The strain number. The institution or collection’s unique internal reference number for the specimen. Used by AMRiN and ALA to link records within the “Darwin Core Archive” file format. For example “CBS 14”.	Free text Value must be present in Core file

	MCL term: http://www.straininfo.net/ns/mcl/2.0/strainNumber DWC term: http://rs.tdwg.org/dwc/terms/catalogNumber	
geneticLocus	Location of a gene or DNA sequence on a chromosome. For example: COX1, TEF, Beta tubulin, ITS, 16S or 18S, 23S or 26S or 28S	Free text
sequenceOrigin	Whether the sequence is of DNA or protein or RNA	Values:“DNA”, “Protein”, “RNA”, null
sequence	Correctly edited final sequence.	Free text
sequenceURL	URL for the sequence that is registered at an authorised registry such as GenBank, BOLD or SwissPROT. DWC term: http://rs.tdwg.org/dwc/terms/associatedSequences	URL
sequenceForwardPrimer	Name of the oligonucleotide used as a primer for the ‘plus’ strand during amplification, or url for its description. For example, “Universal Forward 20mer 5' GTTGTAAAACGACGGCCAGT 3'”	Free text
sequenceReversePrimer	Name of the oligonucleotide used as a primer for the ‘minus’ strand during amplification, or url for its description. For example, “Universal Reverse 20mer 5' CACAGGAAACAGCTATGACC 3'”	Free text

2.5 Measurement File Layout

This class supports sharing of observable or testable characteristics other than temperature and oxygen relationship (which are covered in the core file). Each record in this file identifies a microorganism characteristic and result. Each record should contain a complete tuple.

For example:

- “TL305108”,“Antibiotic Response - cft”,“Resistant”
- “TL305108”,“Antibiotic Response - cp”,“Sensitive”
- “TL305108”,“Antibiotic Response - su”,“Sensitive”

…or:

- “F766”,“Width”,“10-30µm”
- “F766”,“Length”,“25-60µm”
- “F766”,“Flagella”,“2”
- “F766”,“Chloroplasts”,“Parietal”.

Field name	Description	Data Rules
catalogNumber *(previously strainNumber)*	The strain number. The institution or collection’s unique internal reference number for the specimen. Used by AMRiN and ALA to link records within the “Darwin Core Archive” file format. For example “CBS 14”. MCL term: http://www.straininfo.net/ns/mcl/2.0/strainNumber DWC term: http://rs.tdwg.org/dwc/terms/catalogNumber	Free text Value must be present in Core file
measurementType	Identifies what characteristic or feature the observation is about. The nature of the measurement, fact, characteristic, or assertion. For example, “Antibiotic Response - cft”. DWC term: http://rs.tdwg.org/dwc/terms/measurementType	Free text Controlled vocabulary - new terms and synonyms will be flagged when uploading data
measurementValue	The value of the measurement, fact, characteristic, or assertion. For example, the antibiotic response result “Resistant”. DWC term: http://rs.tdwg.org/dwc/terms/measurementValue	Free text

2.6 Bibliography File Layout

Field name	Description	Data Rules
catalogNumber *(previously strainNumber)*	The strain number. The institution or collection’s unique internal reference number for the specimen. Used by AMRiN and ALA to link records within the “Darwin Core Archive” file format.	Free text Value must be present in Core file

	For example “CBS 14”. MCL term: http://www.straininfo.net/ns/mcl/2.0/strainNumber DWC term: http://rs.tdwg.org/dwc/terms/catalogNumber	
bibliographicCitation	A bibliographic reference for the resource. Recommended practice is to include sufficient bibliographic detail to identify the resource as unambiguously as possible. MLC term: http://dublincore.org/documents/dcmi-terms/#terms-bibliographicCitation DWC term: http://purl.org/dc/terms/bibliographicCitation	Free text
title	A name given to the resource. Typically, a Title will be a name by which the resource is formally known. PRISM recommends that magazine publishers use this for the headline of an article. The name of the magazine in which the article appears can be provided in the prism:publicationName element. MCL term: http://dublincore.org/documents/dcmi-terms/#terms-title DWC term: http://purl.org/dc/terms/title	Free text
creator	An entity primarily responsible for making the resource. Examples of a Creator include a person, an organization, or a service. Typically, the name of a Creator should be used to indicate the entity. MCL term: http://dublincore.org/documents/dcmi-terms/#terms-creator DWC term: http://purl.org/dc/terms/creator	Free text
associatedMedia *(previously url)*	The url for the article or unit of content. Refer http://www.prismstandard.org/specifications/2.0/PRISM_prism_namespace_2.0.pdf DWC term: http://rs.tdwg.org/dwc/terms/associatedMedia	URL
source *(previously publicationName)*	Title of the magazine, or other publication, in which a resource was/will be published. Refer http://www.prismstandard.org/specifications/2.0/PRISM_prism_namespace_2.0.pdf	Free text

	DWC term: http://purl.org/dc/terms/source	
volume	Indication of the magazine volume. This element is intended to be used in combination with the number element to specify the magazine issue using the common scheme of Volume and Number. The content should not contain “Vol.” or other abbreviations for “Volume”, it should only be the alphanumeric identifier. Refer http://www.prismstandard.org/specifications/2.0/PRISM_prism_namespace_2.0.pdf	Free text
number	Indication of the magazine issue. This element is intended to be used in combination with the volume element to specify the magazine issue using the common scheme of Volume and Number. The content should not contain “No.” or other abbreviations for “Number, it should only be the alphanumeric identifier. Refer http://www.prismstandard.org/specifications/2.0/PRISM_prism_namespace_2.0.pdf	Free text
startingPage	Identifies the first page number for the published print version of the resource. Refer http://www.prismstandard.org/specifications/2.0/PRISM_prism_namespace_2.0.pdf	Free text
pageRange	Identifies the page range for the published print version of the resource. Refer http://www.prismstandard.org/specifications/2.0/PRISM_prism_namespace_2.0.pdf	Free text
modified *(previously issued)*	Date of publication of the resource. Preferred format is YYYY-MM-DDThh:mm:ss±hhmm For example: • “2010-04-01T12:00:00+1000” – April 1 2010, noon AEST • “2010-04-01” – April 1 2010. MCL term: http://dublincore.org/documents/dcmi-terms/#terms-issued	Datetime ISO 8601

	DWC term: http://purl.org/dc/terms/modified	

2.7 Images File Layout

Field name	Description	Data Rules
catalogNumber *(previously strainNumber)*	The strain number. The institution or collection's unique internal reference number for the specimen. Used by AMRiN and ALA to link records within the "Darwin Core Archive" file format. For example "CBS 14". MCL term: http://www.straininfo.net/ns/mcl/2.0/strainNumber DWC term: http://rs.tdwg.org/dwc/terms/catalogNumber	Free text Value must be present in Core file
rightsHolder	The name(s) of the person(s)/organization who owns the copyright of the image or who could give information about the image copyright. For example, "John Doe" or "Dr JR Doe (john.doe@scs.edu.au") DWC term: http://purl.org/dc/terms/rightsHolder	Free text
creator	The name(s) of the person(s)/organization who took the image. An entity primarily responsible for making the resource. For example, "John Doe" or "Dr JR Doe (john.doe@scs.edu.au") DWC term: http://purl.org/dc/terms/creator	Free text
rights	Identifies the terms under which the image is made available, who may use it, and how it may be used. http://creativecommons.org/choose/ has information about "Creative Commons' licencing, and assistance to select the correct licence. For example, "CC 3.0 by-nc-sa" If not provided, assumed to be "CC 3.0 by-nc-sa". DWC term: http://purl.org/dc/terms/rights	Free text
description	Brief description of the file contents.	Free text

	For example, “Microscopy image of algae”. DWC term: http://purl.org/dc/terms/description	
identifier	Filename of related file that is present in the zip-file archive, or URL for a related image that is not present in the archive, but accessible by internet. DWC term: http://purl.org/dc/terms/identifier	URL

2.8 Recommended Minimum Data

The recommended minimum data supports sharing of “what, when and where” information about specimens. It is a reduced set of the ‘core’ elements, and can be transferred in a single file consistent with the “Simple Darwin Core” standard (see http://darwincore.googlecode.com/svn/trunk/terms/simple/index.htm for more information.

The field names are:

- institutionCode
- collectionCode
- catalogNumber
- scientificName
- basisOfRecord
- otherCatalogNumbers
- eventDate
- sampleOccurrenceType
- sampleSource
- sampleSpecificSite
- country
- stateProvince
- verbatimLatitude
- verbatimLongitude
- coordinateUncertaintyInMeters
- coordinatePrecision
- geodeticDatum