Atlas of Living Australia

BioloMICS Implementation Proposal

Author(s): Bryan Kalms

Version: v1.0

Date: 30 March 2010

File: BioloMICS implementation proposal v1.0.doc

Revision history

Version	Date	Author(s)	Change description
0.1	March 2010	Bryan Kalms	Initial draft
0.2	March 2010	Bryan Kalms	Includes comments from John Tann
1.0	30 March 2010	Bryan Kalms	Includes comments from John Tann. Added offer to build common data model

Table of contents

Introduction

Purpose

The purpose of this document is to offer a way ahead for managing microorganism collections and sharing of information about those collections. In particular, it proposes:

- developing a portal (AMRiN—Australian Microorganism Research information Network) to facilitate information sharing
- providing BioloMICS software to members of the Council of Heads of Australian Collections of Microorganisms (CHACM) who wish to use this software to manage their microorganism collections
- assisting CHACM develop a common data model for microorganisms
- assisting BioloMICS users install and set up the software, including migrating existing data into the new application and establishing protocols and mechanisms to share that data through the ALA and AMRiN.

Background

The ALA has been consulting with the microorganism community to identify ways and means by which the ALA might support the community to share its data, and in particular, share its data with the ALA.

Most recently the ALA hosted a workshop with the community to gain a consensus view of the desired ways and means to manage microorganism collections and share data. At this workshop, the community—representing around 2 dozen organisations—agreed:

- to share their data with each other and the ALA, to the extent permitted by organisational policy and legislation
- BioloMICS was the preferred software solution for managing microorganism collections
- to establish AMRiN as a community tool for sharing information on microorganism activities as well as data on the microorganisms held in collections
- to develop a standard set of metadata to be used to describe microorganism collections
- to consider a proposal from the ALA to support these decisions.

Having considered the implications of these decisions the ALA proposes the way ahead described in this paper.

The proposed way ahead

BioloMICS

Microorganism collections are currently managed using a number of tools, including paper, Excel spreadsheets, Oracle databases and BioloMICS. Not all collection curators have the necessary skills or resources to adopt, and then migrate to, another collection management tool. Accordingly, ALA proposes to assist curators/organisations by:

- providing up to 4 BioloMICS licences per organisation, though 2 licences would be the normal expectation, including paying annual licence fees through to at least the end of June 2012
- liaising and cooperating with the IT departments of participating organisations on the installation of BioloMICS
- if necessary, installing, or assisting in installing, and setting up BioloMICS on hardware provided by the organisation. Note that ALA is not currently able to provide BioloMICS as a hosted service, so all installations must be on a desktop PC or network server. Minimum hardware requirements will be confirmed
- assisting the curator map existing databases and data sets into the AMRiN data model
- assisting in migrating existing databases and data sets to BioloMICS. Any data entry would be the responsibility of the organisation
- liaising with relevant parts of an organisation to facilitate the transfer of data from operational information systems, eg patient care systems, to BioloMICS
- providing initial training in BioloMICS for up to 2 users. This is expected to be of 2 days duration
- advising curators on establishing a web presence for their microorganism collection using BioloMICS
- designing and implementing a suitable mechanism for mobilising data from BioloMICS to both ALA and hence to AMRiN
- providing help desk and technical support for BioloMICS until the end of June 2012.

By default, BioloMICS will be installed with a database configured to the AMRiN data model which, in turn, is based on the CABRI data model—see AMRiN data model below. BioloMICS will also be configured to export data to the ALA/AMRiN, the extent and timing of such sharing being under the control of the curator.

Non-BioloMICS solutions

Where a curator/organisation chooses not to adopt BioloMICS, ALA will work with the curator to 'mobilise' their data for sharing. At this time mobilisation is expected to involve:

- liaising and cooperating with the organisation's IT department to agree the mobilisation solution
- designing a means for regularly providing data upload to ALA/AMRiN, eg development of scripts to export CSV files
- assisting the curator map existing databases and data sets into the AMRiN data model to facilitate data processing by the ALA

- training staff in using the mobilisation solution.

AMRiN

AMRiN is to become the main vehicle for sharing detailed information about microorganism collections. ALA has previously written to the CHACM proposing a number of options for establishing AMRiN. Our understanding is that CHACM has agreed in principle to the option that creates AMRiN as a virtual collection using data held in the ALA. This model is illustrated in Figure 1.

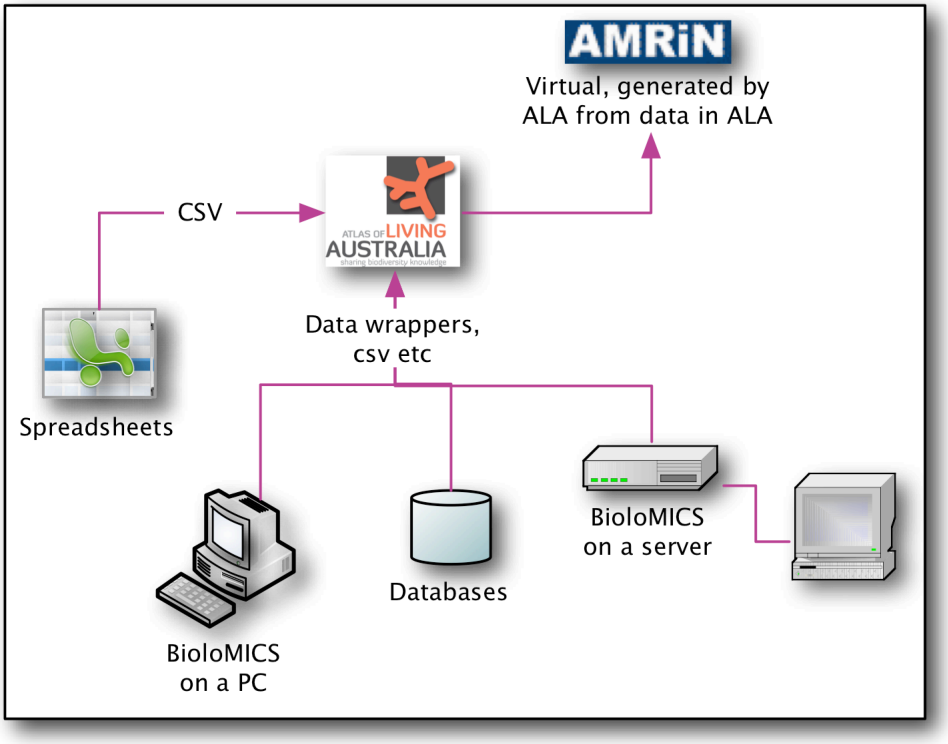

Figure 1 AMRiN as a virtual collection

Accordingly, ALA will work with CHACM and its members to define and deliver the functionality required for AMRiN.

Participant consequences

Organisations participating in this program should recognise and accept the consequences of participation:

- Licence fees may need to be borne by each organisation post June-2012, depending on the availability of future funding for the ALA. Our estimate of such fees is around $600 per licence, depending on the number of participants. Included in this fee will be standard support by Bio-Aware, publisher of BioloMICS, and software updates.
- Each participating organisation will be required to enter into a Data Provider Agreement with the ALA as the basis of information sharing. A copy of the agreement is attached.
- Data fields in BioloMICS or other solutions must comply with the AMRiN data model which allows for additional collection-specific data fields—see AMRiN data model below. This model defines the data to be shared.

- If an organisation requires assistance by ALA to implement BioloMICS, a non-BioloMICS solution or data mobilisation, it must provide access to its facilities and technology and facilitate discussions and liaison with the organisation's IT department.
- Data shared with the ALA will be processed through the normal ALA data upload/cleaning facility. As a result of this process and feedback from ALA/AMRiN users, curators may need to update their data to rectify any deficiencies found, eg incorrect species/strain names.

AMRiN data model

To facilitate information sharing a common AMRiN data model will be developed by CHACM, in consultation with the ALA, and will consist of 3 data sets as illustrated in Figure 2:

1. A set of CABRI-based data fields, consisting of mandatory, recommended and full data fields. This set is expected to vary by type of microorganism and will be the minimum set of data collected and published by the ALA/AMRiN. Data will be publicly available unless it is too sensitive for public release.
2. An optional set of additional data fields covering descriptive data about a microorganism or strain of microorganism. This set can contain as many data fields as agreed with the ALA and will be publicly available, subject to sensitivity constraints. The set may also vary by type of microorganism.
3. If appropriate, a second optional set of data fields containing data that can only be accessed by members of AMRiN and/or CHACM. It, too, may vary by type of microorganism.

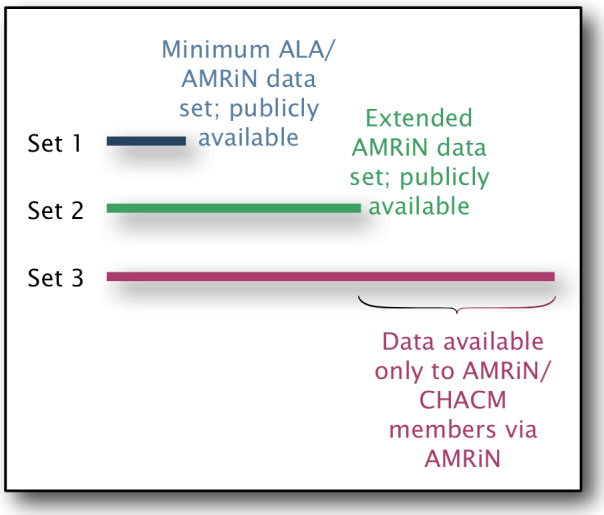

Figure 2 AMRiN data sets

All data sets will be included in the default data base provided as part of the BioloMICS implementation. Data set 1 must be agreed before BioloMICS can be implemented and sets 2 and 3 before AMRiN can be implemented.

ALA is prepared to take the lead in developing this common AMRiN data model.

Timings

An indicative program to implement BioloMICS and establish AMRiN is given in Table 1. A firm program will be available by mid-May 2010.

Table 1 Indicative timings

Activity	Timing	Comments
CHACM formally agrees to create AMRiN as a virtual collection	7 April 2010	Agreement separately sought from CHACM through the hubs options paper
Finalise organisational commitment/ requirements for BioloMICS	30 Apr 2010	Each organisation to respond to this proposal by 30 April 2010
Define AMRiN requirements	3 May–30 Jul 2010	Part of defining the requirements for all hubs
Define AMRiN data model set 1	14 May 2010	
Obtain BioloMICS licences	31 May 2010	
Create default BioloMICS database	18 June 2010	
Agree BioloMICS implementation program	30 June 2010	In consultation with CHACM
Implement BioloMICS	1 July –18 Dec 2010	Up to 24 weeks, assuming 24 participants each taking up to 1 week. Involves: • mapping existing data fields to AMRiN data model • installing BioloMICS software if necessary • migrating data into BioloMICS • liaising and consulting with IT etc
Train BioloMICS users	1 July –18 Dec 2010	
Implement non-BioloMICS solutions	1 July –18 Dec 2010	Allowing up to 1 week per organisation
Harvest data into ALA	1 July –18 Dec 2010	As each organisation is brought 'on line'
Define AMRiN data sets 2 and 3	24 Sept 2010	
Design and build AMRiN	1 Oct 2010– 31 Jan 2011	ALA activity with functionality to be agreed as part of the implementation of the virtual hubs program of work
Customisation of basic AMRiN site	1–15 Feb 2011	CHACM responsibility
Publish AMRiN	16 Feb 2011	Part of ALA Release 7

Actions required

So that the ALA can progress this matter, answers to the following questions should be provided by 30 April 2010 for **each** collection:

1. Do you wish to adopt BioloMICS as a tool for managing your microorganism collection(s)?
2. How many BioloMICS licences do you require?
3. Will you require a desktop PC or a network server installation? ALA is not able to provide a hosted service though BioloMICS data bases can be hosted by the ALA. Do you need hosting of your database?
4. What assistance will you require in implementing BioloMICS? For example, software installation, assistance with data migration, assistance establishing a web presence for your collection using BioloMICS.
5. If you do not wish to adopt BioloMICS, but you wish to share your data with the ALA and AMRiN, what assistance, if any, do you need from the ALA to mobilise your data for sharing?
6. Is there anything else that you consider relevant to adopting BioloMICS or sharing information via the ALA/AMRiN?
7. What specific requirements for AMRiN have you identified so far?

In addition, please advise if you would like ALA to take the lead in developing the common AMRiN data model. If so, please nominate those members of the microorganism community who might be involved in this exercise.