



# **SUMMARY: BioMedBridges resource integration workshop**

1 October 2014, EMBL –EBI, Hinxton

# Table of contents

<b>TABLE OF CONTENTS</b> .....	<b>2</b>
<b>BACKGROUND</b> .....	<b>3</b>
Drivers: main issues for different stakeholders	3
<i>User</i> .....	3
<i>Registries</i> .....	3
<i>Funders</i> .....	3
<b>WORKSHOP AIMS</b> .....	<b>5</b>
Topics for Discussion	5
<b>OUTCOMES</b> .....	<b>6</b>
Resource harmonisation wish list	6
<b>ANNEX 1: BACKGROUND DOCUMENTATION</b> .....	<b>10</b>
<b>ANNEX 2: ATTENDEES</b> .....	<b>11</b>

# Background

Multiple resources are currently being developed and maintained to facilitate and promote information on existing standards, policies, models, identifiers, databases, services, and ontologies. They aim to facilitate end-users looking for data standards/tools/models within a particular topic or applicable to a specific dataset. Integration of these resources together would help users finding comprehensive information about a particular topic and would ensure overall bigger impact for all resources involved.

## Drivers: main issues for different stakeholders

### User

There are several registries with different interfaces and no clear understanding about their differences. Moreover registries which provide complementary functionality are not linking to each other.

### Registries

There is duplication of efforts such as content curation, content definition (schemas and vocabularies) and software maintenance (front end back end).

### Funders

It is very difficult to commit with so many repositories which have a high overlap in content and functionality, especially when these repositories are not working together.

These resources include:

- [BioMedBridges Metadata Model and Mapping registry](#)
- [biosharing.org](http://biosharing.org)
- [identifiers.org](http://identifiers.org)
- [EDAM](#)

— [Tools and Data Services Registry](#)

## Workshop aims

1. Identify common goals shared between these resources
2. Identify areas of duplication and gaps
3. Define a common integration and development strategy

## Topics discussed

- Content overlaps: collaboration & coordination, best strategy for sustainability
- Resource integration (mapping of content as required, linking )
- User interfaces (end-user feedback)
- Use of common ontology
- Define a timeline
- identify blockers for interoperability

# Outcomes

## Resource harmonisation wish list

Task	Owner(s)	Timeframe
Harmonisation of database descriptions used by identifiers.org and biosharing and the tools registry to HCLS <ul style="list-style-type: none"><li>▪ compatibility of service description schema with HCLS subset - maybe</li></ul>	all database registries	
ELIXIR/BMB Tools & Data Services Registry <ul style="list-style-type: none"><li>▪ retire db content (for the databases and simple services, i.e. URIs) from the tools registry</li><li>▪ reimport mandatory content from identifier.org db</li><li>▪ additional content from biosharing.org db</li></ul>	Jon Ison	
Ensure HCLS is adequate for needs of Service Registry, Identifier.org and BioSharing		
Point to Biosharing URL link to the relevant records for the BMB deliverable	BioSharing	
API(s) for all registries to share data <ul style="list-style-type: none"><li>▪ Biosharing</li><li>▪ Identifiers.org</li><li>▪ BMB registry</li><li>▪ Services Registry (need explicit use-case please!)</li></ul>	all registries	

BMB to discuss new name for the standards registry to reflect content	BMB	
Harmonise overlapping DB description between resources	BioSharing Identifiers.org	
Use EDAM format URI in BioSharing	BioSharing	
Xref from BioSharing to EDAM formats terms	BioSharing	
Annotation using EDAM terms for Identifiers.org: data returned by the URI (EDAM data), the format it comes in (EDAM format), the URI parameters (prob. EDAM Data->Identifiers).	Identifiers.org EDAM	
Use a common topic vocabulary to classify resources	All	
BioSharing and EDAM to agree on minimum fields (names, short description, synonyms) that BioSharing will use (via import from EDAM). Agree sharing of curation responsibility respecting existing BioSharing and EDAM curation mechanisms	BioSharing EDAM	
EDAM to drop links to format specifications - these properly the concern of BioSharing (tbc post-discussion with EDAM core developers)	EDAM	
Link from BioSharing to Tools Registry for tools that use (as I/O) a particular format.	BioSharing	
Link from Tools Registry to BioSharing for extended format descriptions and to link (where relevant) to	Jon Ison	

terminologies and checklists		
identify a test set, <ul style="list-style-type: none"> <li>▪ details from each registry to outline data for each</li> </ul>	all	short term
Upload test case MIABIS to Molgenis demo server so that others can see what it looks like	Molgenis dev team	short term
Set up mailing list (Google?) for resources collaboration - we already have bioregistries-discuss - could use it???	Steffi & Jon	
Xrefs between all registries' records	all registries	
Implement user stories for the BMB Service registry	Molgenis dev team	short term
Develop and share high-level development roadmaps from each group (based on existing funding over next year and also informing the development of a proposal.)	everyone	
Harmonise curation / sharing of Identifier regex's between Identifier.org and EDAM	Identifiers.org Jon	
International collaborations (RDA, re3data, NIH, ...)		
community involvement: define carrots/motivation/incentives		
quality metrics and monitoring: informative	ELIXIR metrics Task Force?	
Clarify open governance model and contribution	Jon	



mechanism for EDAM as a high priority		
A general view of where we are heading in a coordinated curation strategy	All	
MoU collaboration	Lead: Jon. Contribute: all	2
Publication	Lead: Susanna. Contribute: all	3
Infradev3 proposal (1-2 pages). Aligned with Rafa's matrix (registries/data types). Deadline 21 October.	Lead: Rafa. Contribute: all	1

## Annex 1: Background documentation

- Mind map produced during the meeting: [.mm \(freeMind\) .pdf](#)
- [Presentations given at the workshop](#):

## Annex 2: Attendees

- Nathalie Conte
- Alejandra Gonzalez-Beltran
- Jon Ison
- Jonathan Jetten
- Rafael Jimenez (Chair)
- Nick Juty
- Fleur Kelpin
- Camille Laibe
- Julie McMurry
- Philippe Rocca-Serra
- Susanna Sansone
- Sarala Wimalaratne
- Tom Hancocks
- Stephanie Suhr
- Helen Parkinson