



A NEAR-COMPLETE REGISTRY OF MOUSE RESOURCES

# INTRODUCTION



## FACTS

- THE MOUSE IS A POPULAR MODEL ORGANISM
- IMMENSE INCREASE OF PRODUCED GENOMIC & PHENOMIC DATA

## RESULT

- WIDE RANGE OF DATA TYPES
- NUMEROUS & SCATTERED ONLINE MOUSE RESOURCES

## PROBLEMS

- DATA RETRIEVAL (WHERE TO LOOK FOR WHAT)
- PRESERVATION & CONSISTENCY OF INFORMATION
- DATA QUALITY & FUTURE RETRIEVAL

# ROLE OF CASIMIR



## CASIMIR IS ABOUT

COORDINATION & SUSTAINABILITY OF INTERNATIONAL MOUSE INFORMATICS RESOURCES

## (WP7) OBJECTIVE

IDENTIFY & LIST RESOURCE CENTERS & MOUSE DATABASES IN EUROPE & WORLD WIDE

## PROBLEMS

- LIST CONSTANTLY UPDATED
- DESCRIPTION OF RESOURCES REQUIRED (TOO) MANY ATTRIBUTES
- DYSFUNCTIONAL EXCEL SHEET

## SOLUTION

- GROUPING OF ATTRIBUTES
- AN EFFICIENT MANAGEMENT SYSTEM

## MOUSE RESOURCE BROWSER IS

AN **ADVANCED** CONTENT MANAGEMENT SYSTEM (**CMS**) FOR MOUSE RESOURCES

## HOSTED ENTITIES

- MOUSE RESOURCES (**220**)
- RESOURCE CATEGORIES (**34**)
- OBO & NON-OBO ONTOLOGIES (**72**)
- MIBBI CHECKLISTS (**21**)
- CASIMIR DEFINED QUALITY CRITERIA (**9**)

## ADVANCED

- WSDL PARSING & WEB SERVICE ANALYSIS ON THE FLY
- WEB SERVICE REMOTE PROCEDURE I-O & COMPLEX CUSTOM TYPE STORING

# ENTITY#1 MOUSE RESOURCE<sup>(1)</sup>



## MOUSE RESOURCE SECTIONS

- **GENERAL**
  - CATEGORIES
  - URL(S)
  - CONTACT INFO
- **ONTOLOGIES & STANDARDS**
  - ONTOLOGIES
  - MINIMUM INFORMATION CHECKLISTS
- **TECHNICAL**
  - WEB SERVICE INFORMATION
  - FILES & DUMPS
  - IMPLEMENTATION DETAILS
- **QUALITY CRITERIA**

## OTHERS ALSO VIEWED

HELPS PROMOTE RELATED RESOURCES

### MUGEN Mouse Database (MMdb)

— List of mouse models for immune processes and immunological diseases.

**General**

Ontologies & Standards

Technical

Quality Criteria

Categories

URI

— Databases

— server online — <http://www.m>

Contact

— C. Chandras (Scientific Curator)  
— V. Aidinis (Scientific Director)

others also viewed

— European Mutant Mouse Archive (EMMA)  
— MyMouse  
— EuroPhenome  
— Institute for Laboratory Animal Research (ILAR)  
— PhenoSITE (Phenotype Semantic Information with Terminology of Experiments)  
— Rodent Phenotyping  
— BioMedNet

# ENTITY#1 MOUSE RESOURCE<sup>(2)</sup>



## WEB SERVICE ACCESS & ANALYSIS

- WSDLs ARE PARSED & ANALYSED ON THE FLY
- ONCE SUCCESSFULLY PARSED STORED IN DB

## WHAT IS STORED

- PROCEDURE NAMES
- PROCEDURE INPUTS & OUTPUTS
- CUSTOM COMPLEX TYPE NAMES
- CUSTOM COMPLEX TYPE PARTICLES

## GOALS

- IDENTIFY COMMON REMOTE PROCEDURES
- STANDARDIZE REMOTE PROCEDURES
  - MINIMUM INFORMATION CHECKLIST
  - ONTOLOGY

## WHAT IS SUPPORTED (SO FAR)

- WSDL 1.1
- WSDL 2.0

### MUGEN Mouse Database (MMdb)

— List of mouse models for immune processes and immunological diseases.

General

Ontologies & Standards

**Technical**

Quality

#### Web Service Access

- [a] — server online — <http://bioit.fleming.gr/mugen/mugenws?WSDL>
- 9 recorded operations
- 23 recorded custom types
- last revision 2008-10-08

#### Dumps & Files

#### others also viewed

- European Mutant Mouse Archive (EMMA)
- MyMouse

# ENTITY#2 CATEGORY



## CATEGORIES

- IMPORTANT FOR GROUPING RESOURCES
- IMPORTANT FOR IDENTIFYING RESOURCES
- A RESOURCE CAN FALL IN MULTIPLE CATEGORIES

### Categories

Anatomy and Histology • Resources 16  
— Update 2008-09-09

Animal Careng • Resources 2  
— Update 2008-09-09

BAC (bacterial artificial chromosome) • Resources 1  
— Update 2008-09-09

Behavior Methods • Resources 4  
— Update 2007-10-31

Bioinformatics • Resources 5  
— Update 2008-07-18

Clinical Methods • Resources 2

## ONTOLOGIES

- INDEX OF 72 ONTOLOGIES
- OBO & NON-OBO ONTOLOGIES ARE INCLUDED
- DEDICATED PAGE FOR EVERY ONTOLOGY WITH:
  - SHORT DESCRIPTION
  - OBO IDENTIFIER
  - LINKS TO FILES & WEB PAGES

### GO — Biological process

— Description — Provides structured controlled vocabularies for the annotation of biological processes, cellular components, and molecular functions.

- Is OBO Ontology
- Non Homemade Ontology

#### Links

- GO Home
- GO OBO File
- GO OWL File



## MINIMUM INFORMATION CHECKLIST

- INDEX OF 21 MINIMUM INFORMATION CHECKLISTS
- DEDICATED PAGE FOR EVERY CHECKLIST WITH:
  - SHORT DESCRIPTION
  - LINKS TO FILES & WEB PAGES

### MIAME

— Description — Minimum Information About a Microarray Experiment

### Links

— [MIAME Home](#)

# ENTITY#5 QUALITY CRITERIA<sup>(1)</sup>



## MATURITY MODELS

- DEFINED BY CASIMIR CONSORTIUM MEMBERS
- IS A LOOSE EVALUATION METHOD
- AVOIDS RANKING OF RESOURCES
- INDICATES COVERAGE ON AREAS OF INTEREST

### Maturity Models

Quality and Consistency

Currency

Accessibility

Output

Technical documentation

Data representation standards

Data structure standards

# ENTITY#5 QUALITY CRITERIA<sup>(2)</sup>



## MATURITY MODELS

- DEFINED BY CASIMIR CONSORTIUM MEMBERS
- IS A LOOSE EVALUATION METHOD
- AVOIDS RANKING OF RESOURCES
- INDICATES COVERAGE ON AREAS OF INTEREST

## (i.e.) QUALITY & CONSISTENCY

- 3 POSSIBLE VALUES
- BEST OPTION IS APPARENT

## HOWEVER

- MM DEPENDS ON NATURE OF RESOURCE
- MM IS OPEN TO INTERPRETATION

### Maturity Models

#### Quality and Consistency

#### Quality and Consistency

- No explicit process for assuring consistency
- Process for assuring consistency with manual curation
- Process for assuring consistency, automatic curation only

# DATA ACQUISITION



## DATA ACQUISITION METHODS

- LITERATURE REVIEW
- GOOGLING
- REVIEW EACH RESOURCE
- ONLINE QUESTIONNAIRE  
<http://bioit.fleming.gr/imouse/>

### i-mouse.org questionnaire

Page 1

Page 2

Page 3

**Page 4**

Page 5

#### Data Structure & vocabularies

##### Compliance with data representation standards

- Data coded by local formalism only
- Some Data coded by recognised controlled vocabulary or ontology or use of MIBBI
- General use of both recognised vocabularies or ontologies, and minimal standards

##### Compliance with data structure standards

- Data structured with local model
- Data structured with formal model (i.e. XML, XML schema)
- Use of recognised standard model (i.e. FUGE)

##### Which MIBBI do or will you use

- CIMR
- IMIAGE
- MAGE-TAB
- MIACA
- MIAME
- MIAME/Env
- MIAME/Null

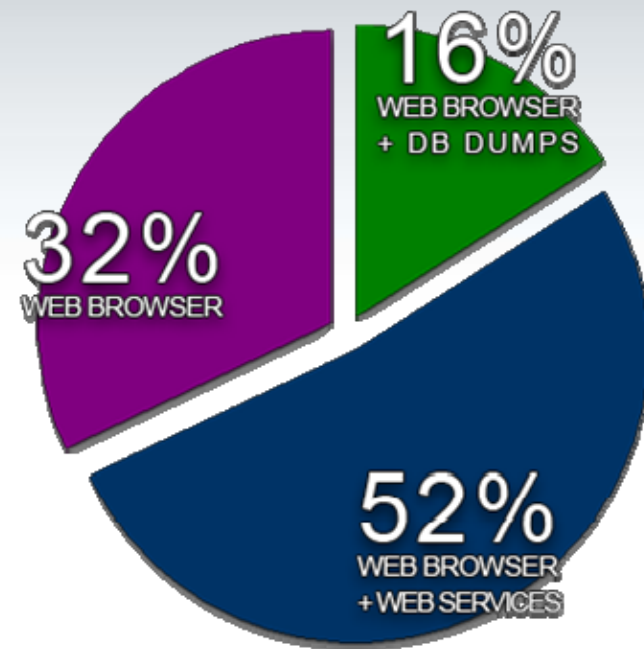
# QUESTIONNAIRE RESPONSES<sup>(1)</sup>



## ACCESS TO DATA

OUT OF 17 REPLIES

- 52% PROVIDE PROGRAMATIC ACCESS METHODS
- 48% DO NOT



# QUESTIONNAIRE RESPONSES<sup>(2)</sup>



## ACCESS TO DATA

OUT OF 17 REPLIES

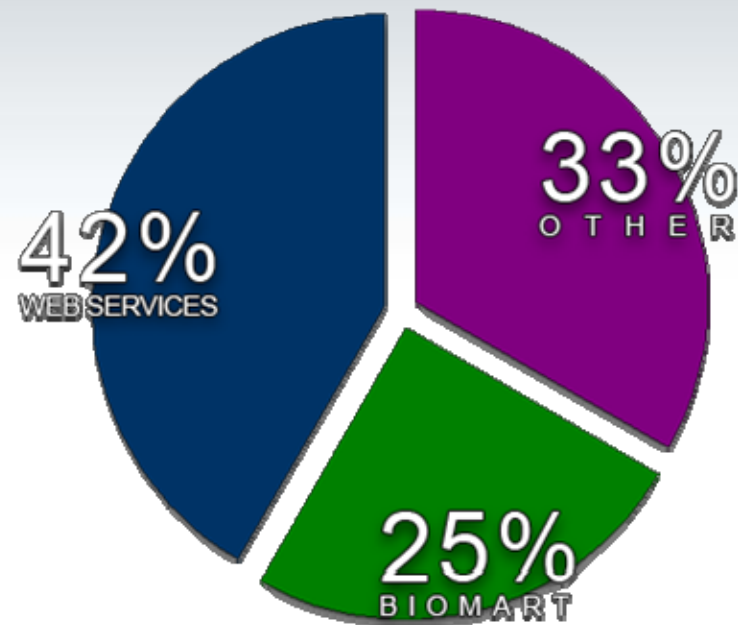
- 52% PROVIDE PROGRAMATIC ACCESS METHODS
- 48% DO NOT

## PREFERRED PA METHODS

- 42% DEVELOPED WEB SERVICES
- 33% WENT FOR OTHER METHODS (PERL MOSTLY)
- 25% OPTED FOR BIOMART

## THE TREND

THERE IS A GROWING TENDENCY AMONGST MOUSE RESOURCES TO DEVELOP WEB SERVICES



# MRB & MOUSE COMMUNITY



## HITS SINCE JUNE 2008

- 7034 HITS
- OVERWHELMING RESPONSE FROM THE MOUSE COMMUNITY
- 25% OF RECORDED RESOURCES ARE PROVIDING CURATORIAL INFORMATION

## GUIDELINES BASED ON LATEST DEVELOPMENTS

- INTEROPERABILITY
- STANDARDISATION

# WORK TO BE DONE



## INTEROPERABILITY

- IMPROVE WSDL PROCESSING & STORING
- ADD **WADL** PROCESSING FEATURES
- INTEGRATE WITH **MOLGENIS** TO PROVIDE A **WEB SERVICE API** FOR MOUSE RESOURCES
- DEVELOP WEB SERVICES FOR MRB

## STANDARDISATION

- COLLECT WEB SERVICE INFORMATION TO COMPILE
- WEB SERVICE ONTOLOGY FOR MOUSE RESOURCES
  - MINIMUM INFORMATION CHECKLIST



# OUTRO



## MRB URL

*<http://www.fleming.gr/mrb>*

## MRB QUESTIONNAIRE URL

*<http://bioit.fleming.gr/imshow>*

## QUESTIONS & SUGGESTIONS???

## ACKNOWLEDGEMENTS

- VASSILIS AIDINIS
- CHRISTINA CHANDRAS
- PAUL SCHOFIELD
- JOHN HANCOCK
- MORRIS SCHWERTZ
- MICHAEL GRUENBERGER
- THE CASIMIR CONSORTIUM