InterMine
Generic object integration and warehousing

FlyMine
An integrated database for *Drosophila* and *Anopheles* genomics

Richard Smith
Genetics Department,
University of Cambridge
InterMine Features

• simple, configurable integration of multiple data sources

• powerful object based query system, not restricted to 'fill in the blanks'
  – web site
  – programming interface
  – query language (OQL-like)

• query optimisation for any model
InterMine Architecture

Web Application -> ObjectStore (O/R) Mapping

ObjectStore (O/R) Mapping -> Relational Database

Relational Database -> QueryOptimiser

QueryOptimiser -> Web service

Web service -> Web Application

Web service <-> SOAP

Queries – Java, IQL

Integration Writer

Integration Writer -> Translation

Translation -> Translation

Configuration:
- merge definition
- primary keys
- source priorities

external data sources: RDBMS, XML, flat files

Master tables

Precomputed tables
Data Model Independent

- auto-generation from XML model definition
- configurable merging of multiple models
# Web Application

## InterMine
Object integration and warehousing software

[Home] [Current query] [New query] [Examples] [Log in] [Help]

Build your query using this page [help...]

<table>
<thead>
<tr>
<th>Model browser</th>
<th>Constraints on the current query</th>
</tr>
</thead>
<tbody>
<tr>
<td>Click on &quot;show&quot; links to add fields to the results table</td>
<td>Click on a class name to view its fields in the left-hand pane</td>
</tr>
<tr>
<td><img src="employee.png" alt="Employee" /> show ▶</td>
<td><img src="employee.png" alt="Employee" /> constrain ▶</td>
</tr>
<tr>
<td><img src="age.png" alt="age" /> int show ▶</td>
<td><img src="age.png" alt="age" /> int constrain ▶</td>
</tr>
<tr>
<td><img src="department.png" alt="department" /> Department show ▶</td>
<td><img src="department.png" alt="department" /> Department constrain ▶</td>
</tr>
<tr>
<td><img src="company.png" alt="company" /> Company show ▶</td>
<td><img src="company.png" alt="company" /> name String constrain ▶</td>
</tr>
<tr>
<td><img src="employees.png" alt="employees" /> Employee collection show ▶</td>
<td><img src="employees.png" alt="employees" /> LIKE finance ▶</td>
</tr>
<tr>
<td><img src="manager.png" alt="manager" /> Manager show ▶</td>
<td><img src="manager.png" alt="manager" /> name String show ▶</td>
</tr>
<tr>
<td><img src="name.png" alt="name" /> String show ▶</td>
<td><img src="name.png" alt="name" /> fullTime boolean show ▶</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fields selected for output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use &quot;&lt;&quot; and &quot;&gt;&quot; to choose the output column order</td>
</tr>
</tbody>
</table>

![Employee](employee.png) [x]
Web Application

- navigate and query any InterMine model
- intuitive interface to create queries
- browse and export results
- save results as a bag – use in other queries and logical operations
- simple definition of 'templates'
- save queries, bags and templates
FlyMine

• *Drosophila* and *Anopheles* data
• extendible core model
• integrates standard formats
  – Sequence Ontology, ensembl, MAGE, PSI, GO
• customised InterMine web application
• future: chado, Uniprot, protein structures
• can be applied to other organisms
InterMine has been in active development for over 18 months. Implementation is in Java with JSP and Struts. All code is available under the LGPL

More information at www.intermine.org

Access the FlyMine Beta at www.flymine.org

The FlyMine team:
Richard Smith, Francois Guillier, Wenyan Ji, Rachel Lyne, Kim Rutherford, Matthew Wakeling, Mark Woodbridge, Andrew Varley, Gos Micklem

FlyMine is funded by the Wellcome Trust (grant no. 067205), awarded to M. Ashburner, G. Micklem, S. Russell, K. Lilley and K. Mizuguchi.