The Learning in Neural Circuits
Research Environment
Managing Living Specimens and Laboratory Data in Islandora

Lingling Jiang  @sprklinginfo
Kim Pham  @tolloid
Kirsta Stapelfeldt  @kstapelfeldt
Blake Richards
Digital Scholarship Unit  @digitalutsc
Learning in Neural Circuits

- Dr. Blake Richards, Department of biological sciences
- New lab set up with undergraduate and graduate students
- Experiments with mice
### VIRAL VECTOR INFUSION SURGERY SHEET

**Injection Coordinates #1:** Brain area: mPFC
- AP: +1.9
- ML: ±0.5
- DV: −1.8

**Injection Coordinates #2:** Brain area: Nucleus accumbens
- AP: −1.1
- ML: 0.24
- DV: −4.3

<table>
<thead>
<tr>
<th>Date</th>
<th>Cage Number</th>
<th>ID</th>
<th>Start</th>
<th>Finish</th>
<th>Weight (g)</th>
<th>Ketoprofen (k.c.)</th>
<th>Pre</th>
<th>Post</th>
<th>AAV used</th>
<th>Drip</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>2/12/15</td>
<td>17</td>
<td>1</td>
<td>8:35</td>
<td>8:35</td>
<td>26.5g</td>
<td>2.65</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/12/15</td>
<td>17</td>
<td>2</td>
<td>4:36</td>
<td>4:36</td>
<td>25.9g</td>
<td>3.56</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/12/15</td>
<td>17</td>
<td>3</td>
<td>6:53</td>
<td>6:53</td>
<td>25.6g</td>
<td>5.53</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/12/15</td>
<td>2</td>
<td>4</td>
<td>6:55</td>
<td>8:30</td>
<td>16.5g</td>
<td>6:52</td>
<td>7:05</td>
<td></td>
<td></td>
<td></td>
<td>Animal died on table</td>
</tr>
</tbody>
</table>

**Isofluorane**

<table>
<thead>
<tr>
<th>Date</th>
<th>Cage Number</th>
<th>ID</th>
<th>Start</th>
<th>Finish</th>
<th>Weight (g)</th>
<th>Ketoprofen (k.c.)</th>
<th>Pre</th>
<th>Post</th>
<th>AAV used</th>
<th>Drip</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>2/12/15</td>
<td>17</td>
<td>1</td>
<td>8:35</td>
<td>8:35</td>
<td>26.5g</td>
<td>2.65</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Animal died on table</td>
</tr>
<tr>
<td>2/12/15</td>
<td>17</td>
<td>2</td>
<td>4:36</td>
<td>4:36</td>
<td>25.9g</td>
<td>3.56</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Animal died on table</td>
</tr>
<tr>
<td>2/12/15</td>
<td>17</td>
<td>3</td>
<td>6:53</td>
<td>6:53</td>
<td>25.6g</td>
<td>5.53</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Animal died on table</td>
</tr>
</tbody>
</table>

**Emia**

<table>
<thead>
<tr>
<th>Date</th>
<th>Cage Number</th>
<th>ID</th>
<th>Start</th>
<th>Finish</th>
<th>Weight (g)</th>
<th>Ketoprofen (k.c.)</th>
<th>Pre</th>
<th>Post</th>
<th>AAV used</th>
<th>Drip</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>2/12/15</td>
<td>17</td>
<td>1</td>
<td>8:35</td>
<td>8:35</td>
<td>26.5g</td>
<td>2.65</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/12/15</td>
<td>17</td>
<td>2</td>
<td>4:36</td>
<td>4:36</td>
<td>25.9g</td>
<td>3.56</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/12/15</td>
<td>17</td>
<td>3</td>
<td>6:53</td>
<td>6:53</td>
<td>25.6g</td>
<td>5.53</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/12/15</td>
<td>2</td>
<td>4</td>
<td>6:55</td>
<td>8:30</td>
<td>16.5g</td>
<td>6:52</td>
<td>7:05</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Notes:**
- Animal died on table
- Animal sacrificed
Data wants to be loved, to feel safe and warm
Requirements

- Accountability to ethics protocols
- Tracking diverse data outputs
- Tracking mice family trees
- Persistent links to data, github repos, resulting images/visualizations and publications
Some example queries for the database

- What is the genotype of female $X$?
- What was the number of pups born for each pregnancy of female $X$?
- How many successful matings did male $Y$ have?
- What protocols did experimental mouse $Z$ undergo?
- Are there any tissue samples for experimental mouse $Z$?
- What mouse did patch-clamp recording $xxxxxxxxx.abf$ come from?
- What are all the data files associated with Figure 1a from paper DOI XX.XXXX?
What is an appropriate data model for tracking living specimens, their offspring, experiments, data, and ethical protocols?
the scheming begins...
<table>
<thead>
<tr>
<th>A</th>
<th>Name</th>
<th>Type</th>
<th>Subject</th>
<th>Classes</th>
<th>Description</th>
<th>Link</th>
<th>Note1</th>
<th>Note2</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Darwin Core</td>
<td>xml schema metadata standard</td>
<td>eeb, biodiversity</td>
<td>biological specimen data</td>
<td><a href="https://wiki.tdwg.org">https://wiki.tdwg.org</a></td>
<td>good for categorizing samples, not elements for biodiversity studies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Gene Ontology (GO)</td>
<td>ontology</td>
<td>genetics, csb</td>
<td>45,000+</td>
<td>the ontology for genomics</td>
<td><a href="http://www.test.ge">http://www.test.ge</a></td>
<td>u of toronto usage, ubiquity, robust, well managed</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>NIFSTD</td>
<td>ontology, db</td>
<td>neuroscience</td>
<td>109,000, NIF Cell has 2800</td>
<td><a href="http://biochemistry.utoronto.ca">http://biochemistry.utoronto.ca</a></td>
<td>comprehensive, apparently can be imported in GO</td>
<td>open source?</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>OBO</td>
<td>foundry</td>
<td>biology</td>
<td>n/a</td>
<td>.obo format for lots of bio/medical ontologies</td>
<td><a href="http://www.obofou">http://www.obofou</a></td>
<td>third preferred format for moving data between ontologies</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>ISA-Tab</td>
<td>metadata standard but it's tab delimited not xsd</td>
<td>general bio</td>
<td>n/a</td>
<td>tab format, not xml schema, not owl or rdf but can be converted AFTER the fact</td>
<td><a href="http://www.dcc.ac">http://www.dcc.ac</a></td>
<td>dataverse uses this, harvard linc - simple, connect to binary data, rdf, nature publishing group</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>RDF</td>
<td>metadata standard</td>
<td>general</td>
<td>n/a</td>
<td>metadata for modelling semantic data</td>
<td><a href="http://www.dcc.ac">http://www.dcc.ac</a></td>
<td>required for owl ontologies - preferred format for ontologies</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>OWL</td>
<td>ontology</td>
<td>general</td>
<td>n/a</td>
<td>format to capture semantics in RDF</td>
<td><a href="http://www.dcc.ac">http://www.dcc.ac</a></td>
<td>standard spec for ontologies - second preferred format</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>IMSR</td>
<td>db</td>
<td>mouse db</td>
<td>n/a</td>
<td>find mouse strains</td>
<td><a href="http://www.findmice.org">http://www.findmice.org</a></td>
<td>specific to area of study</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>JAX</td>
<td>db</td>
<td>mouse db</td>
<td>n/a</td>
<td>another mouse resource, find by genotype/phenotype etc</td>
<td><a href="http://phenome.jax.org">http://phenome.jax.org</a></td>
<td>specific to area of study</td>
<td></td>
</tr>
</tbody>
</table>
Darwin Core, but Not Simple

- Simple Darwin Core was insufficient to capture all of the metadata
- Used full Darwin Core instead
Content model for each -> For now metadata only (DWC datastream)
Darwin Core Relationships

Relationships are stored in Data Session:

**DataSession: Experiment**

```xml
<dwc:ResourceRelationship>
  <dwc:relatedResourceID>USER INPUT</dwc:relatedResourceID>
  <dwc:relationshipOfResource>isSessionOf</dwc:relationshipOfResource>
</dwc:ResourceRelationship>
```

**DataSession: Mouse**

```xml
<dwc:ResourceRelationship>
  <dwc:resourceID>USER INPUT</dwc:resourceID>
  <dwc:relationshipRemarks>USER INPUT</dwc:relationshipRemarks>
  <dwc:relatedResourceID>USER INPUT</dwc:relatedResourceID>
</dwc:ResourceRelationship>
```

Relationships are stored in Experiment:

**Experiment: Protocol**

```xml
<dwc:ResourceRelationship>
  <dwc:relatedResourceID>43</dwc:relatedResourceID>
  <dwc:relationshipOfResource>isGovernedBy</dwc:relationshipOfResource>
</dwc:ResourceRelationship>
```

Relationships are stored in Mouse:

```xml
<dwc:ResourceRelationship>
  <dwc:relatedResourceID/>
  <dwc:relationshipOfResource>hasMother</dwc:relationshipOfResource>
</dwc:ResourceRelationship>
<dwc:ResourceRelationship>
  <dwc:relatedResourceID/>
  <dwc:relationshipOfResource>hasFather</dwc:relationshipOfResource>
</dwc:ResourceRelationship>
```

when there is an arrow pointing away from an object, this means that the relationship is stored in that object.

a mouse can have 0..1 protocols, a protocol can have zero to many mice. nothing to do with # of relationships an object records
Mouse Form

Individual Mouse Information

Mouse Identifier* 

Mouse Type 
Not yet defined, Experiment, Breeding

Strain 
BALB/cJ, 129/Sv, C57BL/6J, NULL

Genotype 
Wild-type, PV-IREs-Cre heterozygous, PV-IREs

Gender 
Not yet defined, female, male

Cage Information

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>(+)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cage ID</td>
<td>Text</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cage Status</td>
<td>born</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Date of Entry</td>
<td>Text</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Add</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Purchased 

Purchased From 

Mouse Form
Add Mouse directly when:
- new parent instance
- purchased mice
- mice with no birth record

- Required fields: mouse identifier

Mouse type
NULL, Not yet defined, Experiment, Breeding

Strain
NULL, BALB/cJ, 129/Sv, C57BL/6J

Genotype
NULL, Wild-type, PV-IREs-Cre heterozygous, PV-IREs-Cre homozygous, VGAT-ChR2-YFP heterozygous, VGAT-ChR2-YFP homozygous, SST-ines-cne homozygous

Gender
NULL, Not yet defined, female, male

NOTE: Not yet defined options need to be searchable

NO PROGRAMMING IS REQUIRED
# Data Session Form

<table>
<thead>
<tr>
<th>Data Session ID*</th>
<th>Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affiliated Experiment ID*</td>
<td>autocomplete</td>
</tr>
<tr>
<td>Collection date</td>
<td>Text</td>
</tr>
<tr>
<td>Experimenter Name</td>
<td>Text</td>
</tr>
<tr>
<td>Description</td>
<td>Text</td>
</tr>
</tbody>
</table>

## Associated Data Files

<table>
<thead>
<tr>
<th>File1</th>
<th>File2</th>
<th>File3</th>
<th>File4</th>
<th>+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Filename</td>
<td>Text</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FileType</td>
<td>Options</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Notes</td>
<td>Text</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mouse ID</td>
<td>autocomplete</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Birth of a Magic Form

- Requirements call for magic - islandora forms doesn’t make sense for their workflow
- use of Drupal forms
### Islandora Repository

#### Birth Identifier *
- 0028

#### Cage born ID
- BT-0001 (linc:580)

#### Mother *
- 0-0-0022-03 (linc:772)

#### Father *
- 0-0-0019-02 (linc:768)
- 0-0-0022-03 (linc:772)
- 0-0-0009-02 (linc:726)
- 0-0-0009-04 (linc:728)
- **0-0-0017-06 (linc:765)**
- 0-0-0015-04 (linc:757)
- 0-0-0013-07 (linc:751)
- 0-0-0015-05 (linc:758)
- 0-0-0015-03 (linc:756)
Wean Date
05/04/2015

Set Up Date
05/26/2015

Children Information

X segment in S-X-III-NN ID schema *
- wild-type

Total number of children born *
- 3

Child #1 ID
- 0-0-0028-01

Child #2 ID
- 0-0-0028-02

Child #3 ID
- 0-0-0028-03
Repository / Mice / 0-0-0001-02

In collections

- Mice

Details

- Last Modified: 2015-03-03T01:37:02.503Z
- Identifier: 0-0-0001-02
- Identifier DC: 0-0-0001-02
- Sex: female
- Father: 1-0-0001-02
Islandora Solr Views!
Repository / Mice / 0-0-0001-02

0-0-0001-02

View Manage

In collections

- Mice

Details

Last Modified: 2015-03-03T01:37:02.503Z
Identifier: 0-0-0001-02
Identifier DC: 0-0-0001-02
Sex: female

Related Objects

Father: 1-0-0001-02
Mother: 1-0-0001-01
Current Cage: B-0007

Children

0-0-0009-02
0-0-0009-04
0-0-0017-06
0-0-0015-04
0-0-0015-05
0-0-0015-03
0-0-0015-06
There are security updates available for one or more of your modules or themes. To ensure the security of your server, you should update immediately! See the available updates page for more information and to install your missing updates.

For This block (override)

Type: string
This display does not have a source for contextual filters, so no contextual filter value will be available unless you select 'Provide default'.

WHEN THE FILTER VALUE IS NOT AVAILABLE

- Display all results for the specified field
- Provide default value

Type

Path component

The numbering starts from 1, e.g. on the page admin/structure/types, the 3rd path component is "types".

- Use path alias
  - Use path alias instead of internal path.

- Hide view
- Display a summary

Apply (this display) Cancel Remove
Reverse Relationship (Not in Object)
Display name: Block

Title: Associated Mice

Format: Unformatted list

Show: Fields

Fields:
- Islandora Solr: dwc.Livingcage_location_PID (dwc.Livingcage_location_PID)
- Islandora Solr: dwc.Relationship_hasFather_PID (dwc.Relationship_hasFather_PID)
- Islandora Solr: PID (PID)
- Islandora Solr: dwc.Relationship_hasMother_PID (dwc.Relationship_hasMother_PID)
- Islandora Solr: dwc.currentCage_location_PID (dwc.currentCage_location_PID)
- Islandora Solr: dwc.Relationship_isDataMeasurementOf_PID (dwc.Relationship_isDataMeasurementOf_PID)
- Islandora Solr: dwc.Relationship_is GovernedBy_PID (dwc.Relationship_is GovernedBy_PID)

Block name: None
Access: None
Use pager: Mini
More link: No

Contextual Filters:
- Islandora Solr: dwc.currentCage_location_PID

Relationships:

No Results Behavior:

Exposed Form:
Exposed form style: Basic

Other:
Machine Name: block
Comment: No comment
Use AJAX: No
Hide attachments in summary: No
Hide contextual links: No
Query settings: Settings
Field Language: Current user's language
Caching: None
Link display: Page
WHEN THE FILTER VALUE IS NOT AVAILABLE

- Display all results for the specified field
- Provide default value

**Type**

PHP Code

```
$path = current_path();
$part = explode('/', $path);
return '*' . $part[2]. '*';
```

Enter PHP code that returns a value to use for this filter. Do not use <?php ?>. You must return only a single value for just this filter. Some variables are available: the view object.
Challenges - If we knew then...

- Data structures != Data workflow
- New Lab = = Prone to change
- Living Research Labs don’t always get prime wireless locations
What’s next?

back to scheming...
Contact us

● Lingling Jiang, Applications Developer, University of Toronto Scarborough, @sprklinginfo
● Kim Pham, Digital Projects & Technologies Librarian, University of Toronto Scarborough, @tolloid
● Kirsta Stapelfeldt, Coordinator, Digital Scholarship Unit, University of Toronto Scarborough, @kstapelfeldt

● Digital Scholarship Unit: digitalscholarship@utsc.utoronto.ca
Progression of the Approach

1: Vanilla Islandora (Collections in collections in collections) --> unusable for Blake and his team

2: custom building --> custom content models, forms, & indexing utilizing open standards (such as Darwin Core)
Mice ID Schema

S–X–III–NN where:

- S = 0 in-house bred, 1 for purchased
- X = 0 for wild-type, 1 for transgenic
- IIII = 4-digit litter/shipment ID number
- NN = 2-digit mouse number (numbers are assigned to each mouse in a litter/shipment)

Initial litter #

0028

Initial litter number used to auto-generate Mice ID III segment.

Initial shipment #

0028

Initial shipment number used to auto-generate Mice ID III segment.
Batch Ingest

Nod to UTL’s development of Zip importer that can be used with DWC

Batch ingest DWC files in appropriate relationship order (parents first, then children) - find and replace