owmeta Documentation

Release 0.11.3.dev0

owmeta

Sep 26, 2019

Contents

Our main README is available online on Github.[1] This documentation contains additional materials beyond what is covered there.

Contents:

[1] http://github.com/openworm/owmeta

CHAPTER 1

owmeta

1.1 owmeta package

1.1.1 owmeta

OpenWorm Unified Data Abstract Layer.

An introduction to owmeta can be found in the README on our Github page.

Most statements correspond to some action on the database. Some of these actions may be complex, but intuitively `a.B()`, the Query form, will query against the database for the value or values that are related to `a` through `B`; on the other hand, `a.B(c)`, the Update form, will add a statement to the database that `a` relates to `c` through `B`. For the Update form, a Statement object describing the relationship stated is returned as a side-effect of the update.

The Update form can also be accessed through the set() method of a Property and the Query form through the get() method like:

```
a.B.set(c)
```

and:

```
a.B.get()
```

The get() method also allows for parameterizing the query in ways specific to the Property.

`owmeta.`**`loadConfig`**(*f*)
Load configuration for the module.

`owmeta.`**`loadData`**(*conf*, *data='OpenWormData/WormData.n3'*, *dataFormat='n3'*, *skipIfNewer=False*)
Load data into the underlying database of this library.

XXX: This is only guaranteed to work with the ZODB database.

Parameters

- **`data`** – (Optional) Specify the file to load into the library

- **dataFormat** – (Optional) Specify the file format to load into the library. Currently n3 is supported
- **skipIfNewer** – (Optional) Skips loading of data if the database file is newer than the data to be loaded in. This is determined by the modified time on the main database file compared to the modified time on the data file.

owmeta.**disconnect**(*c=False*)
Close the database.

owmeta.**connect**(*configFile=False*, *conf=None*, *do_logging=False*, *data=False*, *dataFormat='n3'*)
Load desired configuration and open the database

Parameters

- **configFile** – (Optional) The configuration file for owmeta
- **conf** – (Optional) a configuration object for the connection. Takes precedence over configFile
- **do_logging** – (Optional) If true, turn on debug level logging
- **data** – (Optional) specify the file to load into the library
- **dataFormat** – (Optional) file format of *data*. Currently n3 is supported

owmeta.**config**(*key=None*)
Gets the main configuration for the whole owmeta library.

Returns the instance of the Configure class currently operating.

1.1.2 Subpackages

owmeta.commands package

Various commands of the same kind as *OWM*, mostly intended as sub-commands of OWM.

Submodules

owmeta.commands.bundle module

Bundle commands

exception owmeta.commands.bundle.**NoBundleLoader**(*bundle_name*)
Bases: *owmeta.command_util.GenericUserError*

Thrown when a loader can't be found for a loader

class owmeta.commands.bundle.**OWMBundle**(*parent*)
Bases: object

Bundle commands

checkout(*self*, *bundle_name*)
Switch to the named bundle

Parameters

bundle_name [str] Name of the bundle to switch to

deploy(*self*, *bundle_name*, *remotes=None*)
Deploys a bundle to a remote. The target remotes come from project and user settings or, if provided, the parameters

Parameters

bundle_name [str] Name of the bundle to deploy

remotes [str] Names of the remotes to deploy to

deregister(*self*, *bundle_name*)
Remove a bundle from the project

Parameters

bundle_name [str] The name of the bundle to deregister

fetch(*self*, *bundle_name*)
Retrieve a bundle by name, possibly from remotes, and put in the local bundle repository

Parameters

bundle_name [str] The name of the bundle to retrieve

install(*self*, *bundle_name*)
Install the bundle to the local bundle repository for use across projects on the same machine

Parameters

bundle_name [str] Name of the bundle to install

list(*self*)
List registered bundles in the current project.

To list bundles within the local repo or a remote repo, use the `repo query` sub-command.

load(*self*, *input_file_name*)
Load a bundle from a file

Parameters

input_file_name [str] The source file of the bundle

register(*self*, *descriptor*)
Register a bundle within the project

Registering a bundle adds it to project configuration and records where the descriptor file is within the project's working tree. If the descriptor file moves it must be re-registered at the new location.

Parameters

descriptor [str] Descriptor file for the bundle

save(*self*, *bundle_name*, *output*)
Write a bundle to a file

Writing the bundle to a file writes the bundle descriptor, constituent graphs, and attached files to an archive. The bundle can be in the local bundle repository, a remote, or registered in the project.

Parameters

bundle_name [str] The bundle to save

output [str] The target file

owmeta.data_trans package

Data translators

Some `DataSource` and `DataTranslator` types. Some deal with generic file types (e.g., comma-separated values) while others are specific to the format of a kind of file housed in owmeta.

Submodules

owmeta.data_trans.bibtex module

class owmeta.data_trans.bibtex.**BibTexDataSource**(*bibtex_file_name*, ***kwargs*)
Bases: *owmeta.data_trans.local_file_ds.LocalFileDataSource*

File name [DatatypeProperty] Attribute: `file_name`

Torrent file name [DatatypeProperty] Attribute: `torrent_file_name`

MD5 hash [DatatypeProperty] Attribute: `md5`

SHA-256 hash [DatatypeProperty] Attribute: `sha256`

SHA-512 hash [DatatypeProperty] Attribute: `sha512`

Input source [ObjectProperty] Attribute: `source`

The data source that was translated into this one

Translation [ObjectProperty] Attribute: `translation`

Information about the translation process that created this object

Description [DatatypeProperty] Attribute: `description`

Free-text describing the data source

class owmeta.data_trans.bibtex.**BibTexDataTranslator**(***kwargs*)
Bases: *owmeta.datasource.DataTranslator*

Input type(s): *owmeta.data_trans.bibtex.BibTexDataSource*

Output type(s): *owmeta.data_trans.bibtex.EvidenceDataSource*

URI: None

input_type
alias of *BibTexDataSource*

output_type
alias of *EvidenceDataSource*

translate(*data_source*)
Notionally, this method takes a data source, which is translated into some other data source. There doesn't necessarily need to be an input data source.

class owmeta.data_trans.bibtex.**EvidenceDataSource**(**args*, ***kwargs*)
Bases: *owmeta.datasource.DataSource*

Context [ObjectProperty] Attribute: `evidence_context`

The context

Input source [ObjectProperty] Attribute: `source`

The data source that was translated into this one

Translation [ObjectProperty] Attribute: translation

Information about the translation process that created this object

Description [DatatypeProperty] Attribute: description

Free-text describing the data source

owmeta.data_trans.common_data module

owmeta.data_trans.connections module

class owmeta.data_trans.connections.**ConnectomeCSVDataSource**(**args*, ***kwargs*)
Bases: *owmeta.data_trans.csv_ds.CSVDataSource*

A CSV data source whose CSV file describes a neural connectome

Basically, this is just a marker type to indicate what's described in the CSV – there's no consistent schema

class owmeta.data_trans.connections.**NeuronConnectomeCSVTranslation**(***kwargs*)
Bases: *owmeta.datasource.GenericTranslation*

class owmeta.data_trans.connections.**NeuronConnectomeCSVTranslator**(***kwargs*)
Bases: *owmeta.data_trans.csv_ds.CSVDataTranslator*

Input type(s): *owmeta.data_trans.connections.ConnectomeCSVDataSource*, *owmeta.data_trans.data_with_evidence_ds.DataWithEvidenceDataSource*

Output type(s): *owmeta.data_trans.data_with_evidence_ds.DataWithEvidenceDataSource*

URI: http://openworm.org/entities/translators/NeuronConnectomeCSVTranslator

output_type
alias of *owmeta.data_trans.data_with_evidence_ds.DataWithEvidenceDataSource*

translation_type
alias of *NeuronConnectomeCSVTranslation*

make_translation(*self*, *sources*)
It's intended that implementations of DataTranslator will override this method to make custom Translations according with how different arguments to Translate are (or are not) distinguished.

The actual properties of a Translation subclass must be defined within the 'translate' method

translate(*self*, *data_source*, *neurons_source*, *muscles_source*)
Notionally, this method takes a data source, which is translated into some other data source. There doesn't necessarily need to be an input data source.

class owmeta.data_trans.connections.**NeuronConnectomeSynapseClassTranslation**(***kwargs*)
Bases: *owmeta.datasource.GenericTranslation*

class owmeta.data_trans.connections.**NeuronConnectomeSynapseClassTranslator**(***kwargs*)
Bases: *owmeta.data_trans.csv_ds.CSVDataTranslator*

Adds synapse classes to existing connections

output_type
alias of *owmeta.data_trans.data_with_evidence_ds.DataWithEvidenceDataSource*

translation_type
alias of *NeuronConnectomeSynapseClassTranslation*

make_translation(*self*, *sources*)
It's intended that implementations of DataTranslator will override this method to make custom Translations according with how different arguments to Translate are (or are not) distinguished.

The actual properties of a Translation subclass must be defined within the 'translate' method

translate(*self*, *data_source*, *neurotransmitter_source*)
Notionally, this method takes a data source, which is translated into some other data source. There doesn't necessarily need to be an input data source.

owmeta.data_trans.context_datasource module

class owmeta.data_trans.context_datasource.**VariableIdentifierContext**(*maker=None*, ***kwargs*)
Bases: owmeta.data_trans.context_datasource.VariableIdentifierMixin, *owmeta.context.Context*

A Context that gets its identifier and its configuration from its 'maker' passed in at initialization

This is defined so that the __init__ method gets a contextualized instance, allowing for statements made in __init__ to be contextualized.

class owmeta.data_trans.context_datasource.**VariableIdentifierContextDataObject**(*maker=None*, ***kwargs*)
Bases: owmeta.data_trans.context_datasource.VariableIdentifierMixin, *owmeta.contextDataObject.ContextDataObject*

A ContextDataObject that gets its identifier and its configuration from its 'maker' passed in at initialization

defined_augment(*self*)
This fuction must return False if `identifier_augment()` would raise an `IdentifierMissingException`. Override it when defining a non-standard identifier for subclasses of DataObjects.

owmeta.data_trans.context_merge module

class owmeta.data_trans.context_merge.**ContextMergeDataTranslator**(***kwargs*)
Bases: *owmeta.datasource.DataTranslator*

Input type(s): *owmeta.datasource.OneOrMore* (*owmeta.data_trans.data_with_evidence_ds.DataWithEvidenceDataSource*)

Output type(s): *owmeta.data_trans.data_with_evidence_ds.DataWithEvidenceDataSource*

URI: http://openworm.org/entities/translators/ContextMergeDataTranslator

output_type
alias of *owmeta.data_trans.data_with_evidence_ds.DataWithEvidenceDataSource*

translate(*self*, **sources*)
Notionally, this method takes a data source, which is translated into some other data source. There doesn't necessarily need to be an input data source.

owmeta.data_trans.csv_ds module

class owmeta.data_trans.csv_ds.**CSVDataSource**(**args*, ***kwargs*)
Bases: *owmeta.data_trans.local_file_ds.LocalFileDataSource*

CSV file name [DatatypeProperty] Attribute: csv_file_name

Header column names [DatatypeProperty] Attribute: csv_header

CSV field delimiter [DatatypeProperty] Attribute: csv_field_delimiter

File name [DatatypeProperty] Attribute: file_name

Torrent file name [DatatypeProperty] Attribute: torrent_file_name

MD5 hash [DatatypeProperty] Attribute: md5

SHA-256 hash [DatatypeProperty] Attribute: sha256

SHA-512 hash [DatatypeProperty] Attribute: sha512

Input source [ObjectProperty] Attribute: source

The data source that was translated into this one

Translation [ObjectProperty] Attribute: translation

Information about the translation process that created this object

Description [DatatypeProperty] Attribute: description

Free-text describing the data source

class owmeta.data_trans.csv_ds.**CSVDataTranslator**(***kwargs*)
Bases: *owmeta.datasource.DataTranslator*

Input type(s): *owmeta.datasource.DataSource*

Output type(s): *owmeta.datasource.DataSource*

URI: None

class owmeta.data_trans.csv_ds.**CSVHTTPFileDataSource**(***kwargs*)
Bases: *owmeta.data_trans.http_ds.HTTPFileDataSource*

Header column names [DatatypeProperty] Attribute: csv_header

CSV field delimiter [DatatypeProperty] Attribute: csv_field_delimiter

URL [DatatypeProperty] Attribute: url

MD5 hash [DatatypeProperty] Attribute: md5

SHA-256 hash [DatatypeProperty] Attribute: sha256

SHA-512 hash [DatatypeProperty] Attribute: sha512

Input source [ObjectProperty] Attribute: source

The data source that was translated into this one

Translation [ObjectProperty] Attribute: translation

Information about the translation process that created this object

Description [DatatypeProperty] Attribute: description

Free-text describing the data source

owmeta.data_trans.data_with_evidence_ds module

class owmeta.data_trans.data_with_evidence_ds.**DataWithEvidenceDataSource**(**args*, ***kwargs*)

Bases: *owmeta.datasource.DataSource*

Evidence context [ObjectProperty] Attribute: evidence_context

The context in which evidence for the "Data context" is defined

Data context [ObjectProperty] Attribute: data_context

The context in which primary data for this data source is defined

Combined context [ObjectProperty] Attribute: combined_context

Context importing both the data and evidence contexts

Input source [ObjectProperty] Attribute: source

The data source that was translated into this one

Translation [ObjectProperty] Attribute: translation

Information about the translation process that created this object

Description [DatatypeProperty] Attribute: description

Free-text describing the data source

owmeta.data_trans.excel_ds module

class owmeta.data_trans.excel_ds.**XLSXHTTPFileDataSource**(***kwargs*)

Bases: *owmeta.data_trans.http_ds.HTTPFileDataSource*

URL [DatatypeProperty] Attribute: url

MD5 hash [DatatypeProperty] Attribute: md5

SHA-256 hash [DatatypeProperty] Attribute: sha256

SHA-512 hash [DatatypeProperty] Attribute: sha512

Input source [ObjectProperty] Attribute: source

The data source that was translated into this one

Translation [ObjectProperty] Attribute: translation

Information about the translation process that created this object

Description [DatatypeProperty] Attribute: description

Free-text describing the data source

owmeta.data_trans.file_ds module

class owmeta.data_trans.file_ds.**FileDataSource**(***kwargs*)

Bases: *owmeta.datasource.DataSource*

MD5 hash [DatatypeProperty] Attribute: md5

SHA-256 hash [DatatypeProperty] Attribute: sha256

SHA-512 hash [DatatypeProperty] Attribute: sha512

Input source [ObjectProperty] Attribute: source

The data source that was translated into this one

Translation [ObjectProperty] Attribute: translation

Information about the translation process that created this object

Description [DatatypeProperty] Attribute: description

Free-text describing the data source

owmeta.data_trans.http_ds module

class owmeta.data_trans.http_ds.**HTTPFileDataSource**(***kwargs*)

Bases: *owmeta.data_trans.file_ds.FileDataSource*

URL [DatatypeProperty] Attribute: url

MD5 hash [DatatypeProperty] Attribute: md5

SHA-256 hash [DatatypeProperty] Attribute: sha256

SHA-512 hash [DatatypeProperty] Attribute: sha512

Input source [ObjectProperty] Attribute: source

The data source that was translated into this one

Translation [ObjectProperty] Attribute: translation

Information about the translation process that created this object

Description [DatatypeProperty] Attribute: description

Free-text describing the data source

owmeta.data_trans.local_file_ds module

class owmeta.data_trans.local_file_ds.**LocalFileDataSource**(**args*, ***kwargs*)

Bases: owmeta.capability.Capable, *owmeta.data_trans.file_ds.FileDataSource*

File paths should be relative – in general, path names on a given machine are not portable

accept_capability_provider(*self*, *cap*, *provider*)

The Capable should replace any previously accepted provider with the one given.

owmeta.data_trans.neuron_data module

class owmeta.data_trans.neuron_data.**NeuronCSVDataSource**(**args*, ***kwargs*)

Bases: *owmeta.data_trans.csv_ds.CSVDataSource*

BibTeX files [DatatypeProperty] Attribute: bibtex_files

List of BibTeX files that are referenced in the csv file by entry ID

CSV file name [DatatypeProperty] Attribute: csv_file_name

Header column names [DatatypeProperty] Attribute: csv_header

CSV field delimiter [DatatypeProperty] Attribute: csv_field_delimiter

File name [DatatypeProperty] Attribute: file_name

Torrent file name [DatatypeProperty] Attribute: torrent_file_name

MD5 hash [DatatypeProperty] Attribute: md5

SHA-256 hash [DatatypeProperty] Attribute: sha256

SHA-512 hash [DatatypeProperty] Attribute: sha512

Input source [ObjectProperty] Attribute: source

The data source that was translated into this one

Translation [ObjectProperty] Attribute: translation

Information about the translation process that created this object

Description [DatatypeProperty] Attribute: description

Free-text describing the data source

class owmeta.data_trans.neuron_data.**NeuronCSVDataTranslator**(**args*, ***kwargs*)

Bases: *owmeta.data_trans.csv_ds.CSVDataTranslator*

Input type(s): *owmeta.data_trans.neuron_data.NeuronCSVDataSource*

Output type(s): *owmeta.data_trans.data_with_evidence_ds.DataWithEvidenceDataSource*

URI: http://openworm.org/entities/translators/NeuronCSVDataTranslator

input_type

alias of *NeuronCSVDataSource*

output_type

alias of *owmeta.data_trans.data_with_evidence_ds.DataWithEvidenceDataSource*

translate(*self*, *data_source*)

Notionally, this method takes a data source, which is translated into some other data source. There doesn't necessarily need to be an input data source.

owmeta.data_trans.wormatlas module

class owmeta.data_trans.wormatlas.**WormAtlasCellListDataSource**(**args*, ***kwargs*)

Bases: *owmeta.data_trans.csv_ds.CSVDataSource*

CSV file name [DatatypeProperty] Attribute: csv_file_name

Header column names [DatatypeProperty] Attribute: csv_header

CSV field delimiter [DatatypeProperty] Attribute: csv_field_delimiter

File name [DatatypeProperty] Attribute: file_name

Torrent file name [DatatypeProperty] Attribute: torrent_file_name

MD5 hash [DatatypeProperty] Attribute: md5

SHA-256 hash [DatatypeProperty] Attribute: sha256

SHA-512 hash [DatatypeProperty] Attribute: sha512

Input source [ObjectProperty] Attribute: source

The data source that was translated into this one

Translation [ObjectProperty] Attribute: translation

Information about the translation process that created this object

Description [DatatypeProperty] Attribute: description

Free-text describing the data source

class owmeta.data_trans.wormatlas.**WormAtlasCellListDataTranslation**(***kwargs*)
Bases: *owmeta.datasource.GenericTranslation*

defined_augment(*self*)
This fuction must return False if *identifier_augment()* would raise an IdentifierMissingException. Override it when defining a non-standard identifier for subclasses of DataObjects.

identifier_augment(*self*)
Override this method to define an identifier in lieu of one explicity set.

One must also override *defined_augment()* to return True whenever this method could return a valid identifier. IdentifierMissingException should be raised if an identifier cannot be generated by this method.

Raises

IdentifierMissingException

class owmeta.data_trans.wormatlas.**WormAtlasCellListDataTranslator**(***kwargs*)
Bases: *owmeta.data_trans.csv_ds.CSVDataTranslator*

Input type(s): *owmeta.data_trans.wormatlas.WormAtlasCellListDataSource*, *owmeta.data_trans.data_with_evidence_ds.DataWithEvidenceDataSource*

Output type(s): *owmeta.data_trans.data_with_evidence_ds.DataWithEvidenceDataSource*

URI: http://openworm.org/entities/translators/WormAtlasCellListDataTranslator

output_type
alias of *owmeta.data_trans.data_with_evidence_ds.DataWithEvidenceDataSource*

translation_type
alias of *WormAtlasCellListDataTranslation*

make_translation(*self*, *sources*)
It's intended that implementations of DataTranslator will override this method to make custom Translations according with how different arguments to Translate are (or are not) distinguished.

The actual properties of a Translation subclass must be defined within the 'translate' method

translate(*self*, *data_source*, *neurons_source*)
Notionally, this method takes a data source, which is translated into some other data source. There doesn't necessarily need to be an input data source.

owmeta.data_trans.wormbase module

class owmeta.data_trans.wormbase.**MuscleWormBaseCSVTranslator**(***kwargs*)
Bases: *owmeta.data_trans.csv_ds.CSVDataTranslator*

Input type(s): *owmeta.data_trans.wormbase.WormBaseCSVDataSource*

Output type(s): *owmeta.data_trans.data_with_evidence_ds.DataWithEvidenceDataSource*

URI: http://openworm.org/entities/translators/MuscleWormBaseCSVTranslator

input_type
alias of *WormBaseCSVDataSource*

output_type
alias of *owmeta.data_trans.data_with_evidence_ds.DataWithEvidenceDataSource*

translate(*self*, *data_source*)
Upload muscles and the neurons that connect to them

class owmeta.data_trans.wormbase.**NeuronWormBaseCSVTranslator**(***kwargs*)
Bases: *owmeta.data_trans.csv_ds.CSVDataTranslator*

Input type(s): *owmeta.data_trans.wormbase.WormBaseCSVDataSource*

Output type(s): *owmeta.data_trans.data_with_evidence_ds.DataWithEvidenceDataSource*

URI: http://openworm.org/entities/translators/NeuronWormBaseCSVTranslator

input_type
alias of *WormBaseCSVDataSource*

output_type
alias of *owmeta.data_trans.data_with_evidence_ds.DataWithEvidenceDataSource*

translate(*self*, *data_source*)
Notionally, this method takes a data source, which is translated into some other data source. There doesn't necessarily need to be an input data source.

class owmeta.data_trans.wormbase.**WormBaseCSVDataSource**(**args*, ***kwargs*)
Bases: *owmeta.data_trans.csv_ds.CSVDataSource*

CSV file name [DatatypeProperty] Attribute: csv_file_name

Header column names [DatatypeProperty] Attribute: csv_header

CSV field delimiter [DatatypeProperty] Attribute: csv_field_delimiter

File name [DatatypeProperty] Attribute: file_name

Torrent file name [DatatypeProperty] Attribute: torrent_file_name

MD5 hash [DatatypeProperty] Attribute: md5

SHA-256 hash [DatatypeProperty] Attribute: sha256

SHA-512 hash [DatatypeProperty] Attribute: sha512

Input source [ObjectProperty] Attribute: source

The data source that was translated into this one

Translation [ObjectProperty] Attribute: translation

Information about the translation process that created this object

Description [DatatypeProperty] Attribute: description

Free-text describing the data source

class owmeta.data_trans.wormbase.**WormbaseIonChannelCSVDataSource**(**args*, ***kwargs*)

Bases: *owmeta.data_trans.csv_ds.CSVDataSource*

CSV file name [DatatypeProperty] Attribute: csv_file_name

Header column names [DatatypeProperty] Attribute: csv_header

CSV field delimiter [DatatypeProperty] Attribute: csv_field_delimiter

File name [DatatypeProperty] Attribute: file_name

Torrent file name [DatatypeProperty] Attribute: torrent_file_name

MD5 hash [DatatypeProperty] Attribute: md5

SHA-256 hash [DatatypeProperty] Attribute: sha256

SHA-512 hash [DatatypeProperty] Attribute: sha512

Input source [ObjectProperty] Attribute: source

The data source that was translated into this one

Translation [ObjectProperty] Attribute: translation

Information about the translation process that created this object

Description [DatatypeProperty] Attribute: description

Free-text describing the data source

class owmeta.data_trans.wormbase.**WormbaseIonChannelCSVTranslator**(***kwargs*)

Bases: *owmeta.data_trans.csv_ds.CSVDataTranslator*

Input type(s): *owmeta.data_trans.wormbase.WormbaseIonChannelCSVDataSource*

Output type(s): *owmeta.data_trans.data_with_evidence_ds.DataWithEvidenceDataSource*

URI: http://openworm.org/entities/translators/WormbaseIonChannelCSVTranslator

input_type
alias of *WormbaseIonChannelCSVDataSource*

output_type
alias of *owmeta.data_trans.data_with_evidence_ds.DataWithEvidenceDataSource*

translate(*self*, *data_source*)
Notionally, this method takes a data source, which is translated into some other data source. There doesn't necessarily need to be an input data source.

class owmeta.data_trans.wormbase.**WormbaseTextMatchCSVDataSource**(**args*, ***kwargs*)

Bases: *owmeta.data_trans.csv_ds.CSVDataSource*

initial_cell_column [DatatypeProperty] Attribute: initial_cell_column

The index of the first column with a cell name

cell_type [DatatypeProperty] Attribute: cell_type

The type of cell to be produced

CSV file name [DatatypeProperty] Attribute: csv_file_name

Header column names [DatatypeProperty] Attribute: csv_header

CSV field delimiter [DatatypeProperty] Attribute: csv_field_delimiter

File name [DatatypeProperty] Attribute: file_name

Torrent file name [DatatypeProperty] Attribute: torrent_file_name

MD5 hash [DatatypeProperty] Attribute: md5

SHA-256 hash [DatatypeProperty] Attribute: sha256

SHA-512 hash [DatatypeProperty] Attribute: sha512

Input source [ObjectProperty] Attribute: source

The data source that was translated into this one

Translation [ObjectProperty] Attribute: translation

Information about the translation process that created this object

Description [DatatypeProperty] Attribute: description

Free-text describing the data source

class owmeta.data_trans.wormbase.**WormbaseTextMatchCSVTranslator**(***kwargs*)

Bases: *owmeta.data_trans.csv_ds.CSVDataTranslator*

Input type(s): *owmeta.data_trans.wormbase.WormbaseTextMatchCSVDataSource*

Output type(s): *owmeta.data_trans.data_with_evidence_ds.DataWithEvidenceDataSource*

URI: http://openworm.org/entities/translators/WormbaseTextMatchCSVTranslator

input_type
alias of *WormbaseTextMatchCSVDataSource*

output_type
alias of *owmeta.data_trans.data_with_evidence_ds.DataWithEvidenceDataSource*

translate(*self*, *data_source*)
Notionally, this method takes a data source, which is translated into some other data source. There doesn't necessarily need to be an input data source.

1.1.3 Submodules

owmeta.bibtex module

owmeta.bibtex.**bibtex_to_document**(*bibtex_entry*, *context=None*)
Takes a single BibTeX entry and translates it into a Document object

owmeta.bibtex_customizations module

owmeta.bibtex_customizations.**author**(*record*)
Split author field by 'and' into a list of names.

Parameters **record** (*dict*) – the record.

Returns dict – the modified record.

owmeta.bibtex_customizations.**customizations**(*record*)
Use some functions delivered by the library

Parameters **record** – a record

Returns – customized record

owmeta.bibtex_customizations.**doi**(*record*)

Parameters **record** (*dict*) – the record.

Returns dict – the modified record.

owmeta.biology module

class owmeta.biology.**BiologyType**(***kwargs*)
Bases: *owmeta.dataObject.DataObject*

owmeta.bittorrent module

class owmeta.bittorrent.**BitTorrentDataSourceDirLoader**(*base_directory=None*)
Bases: *owmeta.datasource_loader.DataSourceDirLoader*

load(*self*, *data_source*)
Loads the files for the data source

Parameters

data_source [*owmeta.datasource.DataSource*] The data source to load files for

Returns

A path to the loaded resource

owmeta.bundle module

class owmeta.bundle.**BundleLoader**(*base_directory=None*)
Bases: object

Loads a bundle.

load(*self*, *bundle_name*)
Loads a bundle into the given base directory

class owmeta.bundle.**Descriptor**(*name*)
Bases: object

Descriptor for a bundle

classmethod make(*obj*)
Makes a descriptor from the given object.

class owmeta.bundle.**FilesDescriptor**
Bases: object

Descriptor for files

owmeta.capabilities module

class owmeta.capabilities.**FilePathCapability**
Bases: owmeta.capability.Capability

Provides a file path.

owmeta.capability module

Defines 'capabilities', pieces of functionality that an object needs which must be injected.

A given capability can be provided by more than one capability provider, but, for a given set of providers, only one will be bound at a time. Logically, each provider that provides the capability is asked, in a user-provided preference order, whether it can provide the capability for the *specific* object and the first one which can provide the capability is bound to the object.

The core idea is dependency injection: a capability does not modify the object: the object receives the provider and an identifier for the capability provided, but how the object uses the provider is up to the object. This is important because the user of the object should not condition its behavior on the particular capability provider used, although it may know about which capabilities the object has.

Note, that there may be some providers that lose their ability to provide a capability. This loss should be communicated with a 'CannotProvideCapability' exception when the relevant methods are called on the provider. This *may* allow certain operations to be retried with a provider lower on the capability order, *but* a provider that throws CannotProvide may validly be asked if it can provide the capability again – if it *still* cannot provide the capability, it should communicate that when asked.

Providers may keep state between calls to provide a capability, but their correctness must not depend on any ordering of method calls except that, of course, their `__init__` is called first.

exception owmeta.capability.**CannotProvideCapability**(*cap*, *provider*)
Bases: Exception

Thrown by a *provider* when it cannot provide the capability during the object's execution

exception owmeta.capability.**NoProviderAvailable**(*cap*, *receiver=None*)
Bases: Exception

exception owmeta.capability.**NoProviderGiven**(*cap*, *receiver=None*)
Bases: Exception

owmeta.cell module

class owmeta.cell.**Cell**(*name=None*, *lineageName=None*, ***kwargs*)
Bases: *owmeta.biology.BiologyType*

A biological cell.

All cells with the same name are considered to be the same object.

Parameters

name [str] The name of the cell

lineageName [str] The lineageName of the cell

blast(*self*)
Return the blast name.

Example:

```
>>> c = Cell(name="ADAL")
>>> c.blast() # Returns "AB"
```

Note that this isn't a Property. It returns the blast extracted from the ''first" lineageName saved.

defined_augment(*self*)
This fuction must return False if *identifier_augment()* would raise an

IdentifierMissingException. Override it when defining a non-standard identifier for subclasses of DataObjects.

identifier_augment (*self*, **args*, ***kwargs*)
Override this method to define an identifier in lieu of one explicity set.

One must also override *defined_augment()* to return True whenever this method could return a valid identifier. IdentifierMissingException should be raised if an identifier cannot be generated by this method.

Raises

IdentifierMissingException

description
A description of the cell

divisionVolume
The volume of the cell at division

Example:

```
>>> v = Quantity("600","(um)^3")
>>> c = Cell(lineageName="AB plapaaaap")
>>> c.divisionVolume(v)
```

lineageName
The lineageName of the cell Example:

```
>>> c = Cell(name="ADAL")
>>> c.lineageName() # Returns ["AB plapaaaapp"]
```

name
The 'adult' name of the cell typically used by biologists when discussing C. elegans

owmeta.cell_common module

owmeta.channel module

class owmeta.channel.**Channel**(*name=None*, ***kwargs*)
Bases: *owmeta.biology.BiologyType*

A biological ion channel.

Attributes

Models [Property]

defined_augment (*self*)
This fuction must return False if *identifier_augment()* would raise an IdentifierMissingException. Override it when defining a non-standard identifier for subclasses of DataObjects.

identifier_augment (*self*)
Override this method to define an identifier in lieu of one explicity set.

One must also override *defined_augment()* to return True whenever this method could return a valid identifier. IdentifierMissingException should be raised if an identifier cannot be generated by this method.

Raises

IdentifierMissingException

appearsIn
Cell types in which the ion channel has been expressed

description
A description of the ion channel

expression_pattern
A pattern of expression of this cell within an organism

gene_WB_ID
Wormbase ID of the encoding gene

gene_class
Classification of the encoding gene

gene_name
Name of the gene that codes for this ion channel

model
Get experimental models of this ion channel

name
Ion channel's name

neuroml_file
A NeuroML describing a model of this ion channel

proteins
Proteins associated with this channel

subfamily
Ion channel's subfamily

class owmeta.channel.**ExpressionPattern**(*wormbaseid=None*, ***kwargs*)
Bases: *owmeta.biology.BiologyType*

defined_augment(*self*)
This fuction must return False if *identifier_augment()* would raise an IdentifierMissingException. Override it when defining a non-standard identifier for subclasses of DataObjects.

identifier_augment(*self*)
Override this method to define an identifier in lieu of one explicity set.

One must also override *defined_augment()* to return True whenever this method could return a valid identifier. IdentifierMissingException should be raised if an identifier cannot be generated by this method.

Raises

IdentifierMissingException

description
Natural language description of the expression pattern

wormbaseID
Alias to *wormbaseid*

wormbaseURL
The URL for the expression pattern in Wormbase

wormbaseid
The ID for the expression pattern in Wormbase

owmeta.channel_common module

owmeta.channel_common.**CHANNEL_RDF_TYPE = rdflib.term.URIRef('http://openworm.org/entities/C**
Shared RDF type for channels

owmeta.channelworm module

class owmeta.channelworm.**ChannelModel**(*modelType=None*, **args*, ***kwargs*)
Bases: *owmeta.dataObject.DataObject*

A model for an ion channel.

There may be multiple models for a single channel.

Example usage:

```
# Create a ChannelModel
>>> cm = P.ChannelModel()
# Create Evidence object
>>> ev = P.Evidence(author='White et al.', date='1986')
# Assert
>>> ev.asserts(cm)
>>> ev.save()
```

conductance
The conductance of this ion channel. This is the initial value, and should be entered as a Quantity object.

gating
The gating mechanism for this channel ("voltage" or name of ligand(s))

ion
The type of ion this channel selects for

modelType
The type of model employed to describe a channel

class owmeta.channelworm.**HomologyChannelModel**(***kwargs*)
Bases: *owmeta.channelworm.ChannelModel*

class owmeta.channelworm.**PatchClampChannelModel**(***kwargs*)
Bases: *owmeta.channelworm.ChannelModel*

class owmeta.channelworm.**PatchClampExperiment**(***kwargs*)
Bases: *owmeta.experiment.Experiment*

Store experimental conditions for a patch clamp experiment.

Ca_concentration
Calcium concentration

Cl_concentration
Chlorine concentration

blockers
Channel blockers used for this experiment

cell
The cell this experiment was performed on

cell_age
Age of the cell

initial_voltage
Starting voltage of the patch clamp

ion_channel
The ion channel being clamped

membrane_capacitance
Initial membrane capacitance

mutants
Type(s) of mutants being used in this experiment

patch_type
Type of patch clamp being used ('voltage' or 'current')

pipette_solution
Type of solution in the pipette

owmeta.cli module

owmeta.cli.**additional_args**(*parser*)
Add some additional options specific to CLI

owmeta.cli_command_wrapper module

exception owmeta.cli_command_wrapper.**CLIUserError**
Bases: Exception

class owmeta.cli_command_wrapper.**CLIAppendAction**(*mapper*, *key*, *index=-1*, **args*, ***kwargs*)
Bases: *owmeta.cli_command_wrapper.CLIStoreAction*

class owmeta.cli_command_wrapper.**CLIArgMapper**
Bases: object

Stores mappings for arguments and maps them back to the part of the object they come from

runners = None
Mapping from subcommand names to functions which run for them

class owmeta.cli_command_wrapper.**CLIStoreAction**(*mapper*, *key*, *index=-1*, **args*, ***kwargs*)
Bases: argparse.Action

Interacts with the CLIArgMapper

class owmeta.cli_command_wrapper.**CLIStoreTrueAction**(**args*, ***kwargs*)
Bases: *owmeta.cli_command_wrapper.CLIStoreAction*

class owmeta.cli_command_wrapper.**CLISubCommandAction**(*mapper*, **args*, ***kwargs*)
Bases: argparse._SubParsersAction

owmeta.cli_common module

owmeta.cli_hints module

owmeta.collections module

class owmeta.collections.**Bag**(***kwargs*)

Bases: *owmeta.dataObject.DataObject*

A convenience class for working with a collection of objects

Example:

```
v = Bag('unc-13 neurons and muscles')
n = P.Neuron()
m = P.Muscle()
n.receptor('UNC-13')
m.receptor('UNC-13')
for x in n.load():
    v.value(x)
for x in m.load():
    v.value(x)
# Save the group for later use
v.save()
...
# get the list back
u = Bag('unc-13 neurons and muscles')
nm = list(u.value())
```

defined_augment(*self*)

This fuction must return False if *identifier_augment()* would raise an `IdentifierMissingException`. Override it when defining a non-standard identifier for sub-classes of DataObjects.

identifier_augment(*self*)

Override this method to define an identifier in lieu of one explicity set.

One must also override *defined_augment()* to return True whenever this method could return a valid identifier. `IdentifierMissingException` should be raised if an identifier cannot be generated by this method.

Raises

IdentifierMissingException

add

An alias for `value`

group_name

Alias for `name`

name

The name of the group of objects

value

An object in the group

owmeta.command module

exception owmeta.command.**ConfigMissingException**(*key*)
Bases: *owmeta.command_util.GenericUserError*

exception owmeta.command.**InvalidGraphException**
Bases: *owmeta.command_util.GenericUserError*

Thrown when a graph cannot be translated due to formatting errors

exception owmeta.command.**NoConfigFileError**
Bases: *owmeta.command_util.GenericUserError*

exception owmeta.command.**OWMDirMissingException**
Bases: *owmeta.command_util.GenericUserError*

exception owmeta.command.**StatementValidationError**(*statements*)
Bases: *owmeta.command_util.GenericUserError*

exception owmeta.command.**UnreadableGraphException**
Bases: *owmeta.command_util.GenericUserError*

Thrown when a graph cannot be read due to it being missing, the active user lacking permissions, etc.

class owmeta.command.**OWM**
Bases: object

High-level commands for working with owmeta data

add_graph(*self*, *url=None*, *context=None*, *include_imports=True*)
Fetch a graph and add it to the local store.

Parameters

url [str] The URL of the graph to fetch

context [rdflib.term.URIRef] If provided, only this context and, optionally, its imported graphs will be added.

include_imports [bool] If True, imports of the named context will be included. Has no effect if context is None.

clone(*self*, *url=None*, *update_existing_config=False*, *branch=None*)
Clone a data store

Parameters

url [str] URL of the data store to clone

update_existing_config [bool] If True, updates the existing config file to point to the given file for the store configuration

branch [str] Branch to checkout after cloning

commit(*self*, *message*)
Write the graph to the local repository

Parameters

message [str] commit message

context(*self*, *context=None*, *user=False*)
Read or set current target context for the repository

Parameters

context [str] The context to set

user [bool] If set, set the context only for the current user. Has no effect for retrieving the context

diff (*self*)
Show differences between what's in the working context set and what's in the serializations

fetch_graph (*self*, *url*)
Fetch a graph

Parameters

url [str] URL for the graph

git (*self*, **args*)
Runs git commmands in the .owm directory

Parameters

***args** arguments to git

imports_context (*self*, *context=None*, *user=False*)
Read or set current target imports context for the repository

Parameters

context [str] The context to set

user [bool] If set, set the context only for the current user. Has no effect for retrieving the context

init (*self*, *update_existing_config=False*)
Makes a new graph store.

The configuration file will be created if it does not exist. If it *does* exist, the location of the database store will, by default, not be changed in that file

Parameters

update_existing_config [bool] If True, updates the existing config file to point to the given file for the store configuration

list_contexts (*self*)
List contexts

merge (*self*)

push (*self*)

reconstitute (*self*, *data_source*)
Recreate a data source by executing the chain of translators that went into making it.

Parameters

data_source [str] Identifier for the data source to reconstitute

save (*self*, *module*, *provider=None*, *context=None*)
Save the data in the given context

Saves the "mapped" classes declared in a module and saves the objects declared by the "provider" (see the argument's description)

Parameters

module [str] Name of the module housing the provider

provider [str] Name of the provider, a callble that accepts a context object and adds statements to it. Can be a "dotted" name indicating attribute accesses

context [str] The target context

say(*self*, *subject*, *property*, *object*)
Make a statement

Parameters

subject [str] The object which you want to say something about

property [str] The type of statement to make

object [str] The other object you want to say something about

serialize(*self*, *context=None*, *destination=None*, *format='nquads'*, *include_imports=False*, *whole_graph=False*)
Serialize the current data context or the one provided

Parameters

context [str] The context to save

destination [file or str] A file-like object to write the file to or a file name. If not provided, messages the result.

format [str] Serialization format (ex, 'n3', 'nquads')

include_imports [bool] If true, then include contexts imported by the provided context in the result. The default is not to include imported contexts.

whole_graph: bool Serialize all contexts from all graphs (this probably isn't what you want)

tag(*self*)

translate(*self*, *translator*, *output_key=None*, *output_identifier=None*, *data_sources=()*, *named_data_sources=None*)
Do a translation with the named translator and inputs

Parameters

translator [str] Translator identifier

imports_context_ident [str] Identifier for the imports context. All imports go in here

output_key [str] Output key. Used for generating the output's identifier. Exclusive with output_identifier

output_identifier [str] Output identifier. Exclusive with output_key

data_sources [list of str] Input data sources

named_data_sources [dict] Named input data sources

config_file
The config file name

log_level
Log level

owmdir
The base directory for owmeta files. The repository provider's files also go under here

store_name
The file name of the database store

class owmeta.command.**OWMDirDataSourceDirLoader**(*basedir=None*)
Bases: *owmeta.datasource_loader.DataSourceDirLoader*

can_load(*self*, *data_source*)
Returns true if the *DataSource* can be loaded by this loader

Parameters

data_source [*owmeta.datasource.DataSource*] The data source to load files for

load(*self*, *ident*)
Loads the files for the data source

Parameters

data_source [*owmeta.datasource.DataSource*] The data source to load files for

Returns

A path to the loaded resource

class owmeta.command.**OWMSource**(*parent*)
Bases: object

Commands for working with DataSource objects

create(*self*, *kind*, *key*, **args*, ***kwargs*)
Create the source and add it to the graph.

Arguments are determined by the type of the data source

Parameters

kind [rdflib.term.URIRef] The kind of source to create

key [str] The key, a unique name for the source

derivs(*self*, *data_source*)
List data sources derived from the one given

Parameters

data_source [str] The ID of the data source to find derivatives of

list(*self*, *context=None*, *kind=None*, *full=False*)
List known sources

Parameters

kind [str] Only list sources of this kind

context [str] The context to query for sources

full [bool] Whether to (attempt to) shorten the source URIs by using the namespace manager

list_kinds(*self*, *full=False*)
List kinds of sources

Parameters

full [bool] Whether to (attempt to) shorten the source URIs by using the namespace manager

show(*self*, **data_source*)

Parameters

***data_source** [str] The ID of the data source to show

data
Commands for saving and loading data for DataSources

class owmeta.command.**OWMSourceData**(*parent*)
Bases: object

Commands for saving and loading data for DataSources

retrieve(*self*, *source*, *archive='data.tar'*, *archive_type=None*)
Retrieves the data for the source

Parameters

source [str] The source for data

archive [str] The file name of the archive. If this ends with an extension like '.zip', and no archive_type argument is given, then an archive will be created of that type. The archive name will *not* have any extension appended in any case.

archive_type [str] The type of the archive to create.

class owmeta.command.**SaveValidationFailureRecord**(*user_module*, *stack*, *validation_record*)
Bases: *owmeta.command.SaveValidationFailureRecord*

Create new instance of SaveValidationFailureRecord(user_module, stack, validation_record)

class owmeta.command.**UnimportedContextRecord**
Bases: *owmeta.command.UnimportedContextRecord*

Stored when statements include a reference to an object but do not include the context of that object in the callback passed to *OWM.save*. For example, if we had a callback like this:

```
def owm_data(ns):
    ctxA = ns.new_context(ident='http://example.org/just-pizza-stuff')
    ctxB = ns.new_context(ident='http://example.org/stuff-sam-likes')
    sam = ctxB(Person)('sam')
    pizza = ctxA(Thing)('pizza')
    sam.likes(pizza)
```

it would generate this error because `ctxB` does not declare an import for `ctxA`

Create new instance of UnimportedContextRecord(context, node_index, statement)

owmeta.command_util module

exception owmeta.command_util.**GenericUserError**
Bases: Exception

An error which should be reported to the user. Not necessarily an error that is the user's fault

class owmeta.command_util.**IVar**(*default_value=None*, *doc=None*, *value_type=<class 'str'>*, *name=None*)
Bases: object

A descriptor for instance variables amended to provide some attributes like default values, value types, etc.

class owmeta.command_util.**PropertyIVar**(**args*, ***kwargs*)
Bases: *owmeta.command_util.IVar*

owmeta.configure module

exception owmeta.configure.**BadConf**
Bases: Exception

Special exception subclass for alerting the user to a bad configuration

class owmeta.configure.**ConfigValue**
Bases: object

A value to be configured. Base class intended to be subclassed, as its only method is not implemented

class owmeta.configure.**Configure**(***initial_values*)
Bases: object

A simple configuration object. Enables setting and getting key-value pairs

Unlike a dict, Configure objects will execute a function when retrieving values to enable deferred computation of seldom-used configuration values. In addition, entries in a *Configure* can be aliased to one another.

copy(*self*, *other*)
Copy this configuration into a different object.

Parameters **other** – A different configuration object to copy the configuration from this object into

Returns

get(*self*, *pname*, *default=NO_DEFAULT*)
Get some parameter value out by asking for a key. Note that unlike dict, if you don't specify a default, then a KeyError is raised

Parameters

pname [str] they key of the value you want to return.

default [object] The default value to return if there's no entry for pname

Returns

The value corresponding to the key

link(*self*, **names*)
Call link() with the names of configuration values that should always be the same to link them together

classmethod open(*file_name*)
Open a configuration file and read it to build the internal state.

Parameters **file_name** – configuration file encoded as JSON

Returns a Configure object with the configuration taken from the JSON file

class owmeta.configure.**Configureable**(*conf=None*, ***kwargs*)
Bases: object

An object which can accept configuration. A base class intended to be subclassed.

get(*self*, *pname*, *default=None*)
Gets a config value from this *Configureable*'s conf

See also:

Configure.get

class owmeta.configure.**ImmutableConfigure**(***initial_values*)
Bases: *owmeta.configure.Configure*

owmeta.connection module

class owmeta.connection.**Connection**(*pre_cell=None*, *post_cell=None*, *number=None*, *syntype=None*, *synclass=None*, *termination=None*, ***kwargs*)
Bases: *owmeta.biology.BiologyType*

number
The weight of the connection

post_cell
The post-synaptic cell

pre_cell
The pre-synaptic cell

synclass
The kind of Neurotransmitter (if any) sent between *pre_cell* and *post_cell*

syntype
The kind of synaptic connection. 'gapJunction' indicates a gap junction and 'send' a chemical synapse

termination
Where the connection terminates. Inferred from type of post_cell at initialization

owmeta.context module

class owmeta.context.**ClassContext**(*ident=None*, *imported=()*, *mapper=None*, *key=None*, *base_namespace=None*, ***kwargs*)
Bases: *owmeta.context.Context*

This is defined so that the __init__ method gets a contextualized instance, allowing for statements made in __init__ to be contextualized.

class owmeta.context.**ClassContextMeta**
Bases: *owmeta.context.ContextMeta*

class owmeta.context.**Context**(*ident=None*, *imported=()*, *mapper=None*, *key=None*, *base_namespace=None*, ***kwargs*)
Bases: *owmeta.import_contextualizer.ImportContextualizer*, *owmeta.context.ContextualizableDataUserMixin*

A context. Analogous to an RDF context, with some special sauce

This is defined so that the __init__ method gets a contextualized instance, allowing for statements made in __init__ to be contextualized.

save(*self*, *graph=None*, *inline_imports=False*, *autocommit=True*, *saved_contexts=None*)
Alias to save_context

save_context(*self*, *graph=None*, *inline_imports=False*, *autocommit=True*, *saved_contexts=None*)
Save the context to a graph

class owmeta.context.**ContextContextManager**(*ctx*, *to_import*)
Bases: object

The context manager created when Context::__call__ is passed a dict

class owmeta.context.**ContextMeta**
Bases: *owmeta.contextualize.ContextualizableClass*

class owmeta.context.**ContextualizableDataUserMixin**(**args*, ***kwargs*)
Bases: *owmeta.contextualize.Contextualizable*, *owmeta.data.DataUser*

This is defined so that the __init__ method gets a contextualized instance, allowing for statements made in __init__ to be contextualized.

class owmeta.context.**QueryContext**(*graph*, **args*, ***kwargs*)
Bases: *owmeta.context.Context*

This is defined so that the __init__ method gets a contextualized instance, allowing for statements made in __init__ to be contextualized.

owmeta.contextDataObject module

class owmeta.contextDataObject.**ContextDataObject**(***kwargs*)
Bases: *owmeta.dataObject.DataObject*

Represents a context

owmeta.context_common module

owmeta.context_store module

exception owmeta.context_store.**ContextStoreException**
Bases: Exception

owmeta.contextualize module

class owmeta.contextualize.**Contextualizable**(**args*, ***kwargs*)
Bases: owmeta.contextualize.BaseContextualizable

A BaseContextualizable with the addition of a default behavior of setting the context from the class's 'context' attribute. This generally requires that for the metaclass of the Contextualizable that a 'context' data property is defined. For example:

```
>>> class AMeta(ContextualizableClass):
>>>     @property
>>>     def context(self):
>>>         return self.__context
```

```
>>>     @context.setter
>>>     def context(self, ctx):
>>>         self.__context = ctx
```

```
>>> class A(six.with_metaclass(Contextualizable)):
>>>     pass
```

This is defined so that the __init__ method gets a contextualized instance, allowing for statements made in __init__ to be contextualized.

class owmeta.contextualize.**ContextualizableClass**
Bases: type

A super-type for contextualizable classes

owmeta.contextualize.**contextualize_helper**(*context*, *obj*, *noneok=False*)
Does some extra stuff to make access to the type of a ContextualizingProxy work more-or-less like access to the the wrapped object

owmeta.contextualize.**decontextualize_helper**(*obj*)
Removes contexts from a ContextualizingProxy

owmeta.data module

class owmeta.data.**Data**(*conf=None*, ***kwargs*)
Bases: *owmeta.configure.Configure*

Provides configuration for access to the database.

Usually doesn't need to be accessed directly

Parameters

conf [Configure] A Configure object

closeDatabase(*self*)
Close a the configured database

destroy(*self*)
Close a the configured database

init(*self*)
Open the configured database

init_database(*self*)
Open the configured database

classmethod load(*file_name*)
Load a file into a new Data instance storing configuration in a JSON format

classmethod open(*file_name*)
Load a file into a new Data instance storing configuration in a JSON format

classmethod process_config(*config_dict*, ***kwargs*)
Load a file into a new Data instance storing configuration in a JSON format

class owmeta.data.**DataUser**(**args*, ***kwargs*)
Bases: *owmeta.configure.Configureable*

A convenience wrapper for users of the database

Classes which use the database should inherit from DataUser.

add_reference(*self*, *g*, *reference_iri*)
Add a citation to a set of statements in the database

Parameters **triples** – A set of triples to annotate

add_statements(*self*, *graph*)
Add a set of statements to the database. Annotates the addition with uploader name, etc

Parameters **graph** – An iterable of triples

infer(*self*)
Fire FuXi rule engine to infer triples

retract_statements(*self*, *graph*)
Remove a set of statements from the database.

Parameters **graph** – An iterable of triples

class owmeta.data.**RDFSource**(**kwargs*)
Bases: *owmeta.configure.Configureable*, *owmeta.configure.ConfigValue*

Base class for data sources.

Alternative sources should dervie from this class

get(*self*)
Gets a config value from this `Configureable`'s `conf`

See also:

Configure.get

open(*self*)
Called on `owmeta.connect()` to set up and return the rdflib graph. Must be overridden by sub-classes.

class owmeta.data.**SPARQLSource**(**kwargs*)
Bases: *owmeta.data.RDFSource*

Reads from and queries against a remote data store

```
"rdf.source" = "sparql_endpoint"
```

open(*self*)
Called on `owmeta.connect()` to set up and return the rdflib graph. Must be overridden by sub-classes.

class owmeta.data.**SleepyCatSource**(**kwargs*)
Bases: *owmeta.data.RDFSource*

Reads from and queries against a local Sleepycat database

The database can be configured like:

```
"rdf.source" = "Sleepycat"
"rdf.store_conf" = <your database location here>
```

open(*self*)
Called on `owmeta.connect()` to set up and return the rdflib graph. Must be overridden by sub-classes.

class owmeta.data.**DefaultSource**(**kwargs*)
Bases: *owmeta.data.RDFSource*

Reads from and queries against a configured database.

The default configuration.

The database store is configured with:

```
"rdf.source" = "default"
"rdf.store" = <your rdflib store name here>
"rdf.store_conf" = <your rdflib store configuration here>
```

Leaving unconfigured simply gives an in-memory data store.

open(*self*)
Called on `owmeta.connect()` to set up and return the rdflib graph. Must be overridden by sub-classes.

class owmeta.data.**ZODBSource**(**args*, ***kwargs*)
Bases: *owmeta.data.RDFSource*

Reads from and queries against a configured Zope Object Database.

If the configured database does not exist, it is created.

The database store is configured with:

```
"rdf.source" = "ZODB"
"rdf.store_conf" = <location of your ZODB database>
```

Leaving unconfigured simply gives an in-memory data store.

open(*self*)
Called on `owmeta.connect()` to set up and return the rdflib graph. Must be overridden by sub-classes.

class owmeta.data.**SQLiteSource**(**args*, ***kwargs*)
Bases: owmeta.data.SQLSource

class owmeta.data.**MySQLSource**(**args*, ***kwargs*)
Bases: owmeta.data.SQLSource

class owmeta.data.**PostgreSQLSource**(**args*, ***kwargs*)
Bases: owmeta.data.SQLSource

owmeta.dataObject module

class owmeta.dataObject.**BaseDataObject**(***kwargs*)
Bases: owmeta.identifier_mixin._IdMixin, yarom.graphObject.GraphObject, *owmeta.context.ContextualizableDataUserMixin*

An object backed by the database

Attributes

rdf_type [rdflib.term.URIRef] The RDF type URI for objects of this type

rdf_namespace [rdflib.namespace.Namespace] The rdflib namespace (prefix for URIs) for objects from this class

properties [list of *owmeta.simpleProperty.RealSimpleProperty* or *owmeta.pProperty.Property*] Properties belonging to this object

owner_properties [list of *owmeta.simpleProperty.RealSimpleProperty* or *owmeta.pProperty.Property*] Properties belonging to parents of this object

classmethod DatatypeProperty(**args*, ***kwargs*)
Attach a, possibly new, property to this class that has a simple type (string,number,etc) for its values

Parameters

linkName [str] The name of this property.

owner [*owmeta.dataObject.BaseDataObject*] The owner of this property.

classmethod ObjectProperty(**args*, ***kwargs*)
Attach a, possibly new, property to this class that has a complex BaseDataObject for its values

Parameters

linkName [str] The name of this property.

owner [*owmeta.dataObject.BaseDataObject*] The owner of this property.

value_type [type] The type of BaseDataObject for values of this property

classmethod UnionProperty(**args*, ***kwargs*)
Attach a, possibly new, property to this class that has a simple type (string,number,etc) or BaseDataObject for its values

Parameters

linkName [str] The name of this property.

owner [*owmeta.dataObject.BaseDataObject*] The owner of this property.

clear_po_cache(*self*)
Clear the property-object cache for this object.

This cache is maintained by and shared by the properties of this object. It isn't necessary to clear this cache manually unless you modify the RDFLib graph indirectly (e.g., through the store) at runtime.

contextualize_augment(*context*)
For MappedClass, rdf_type and rdf_namespace have special behavior where they can be auto-generated based on the class name and base_namespace. We have to pass through these values to our "proxy" to avoid this behavior

decontextualize(*self*)
Return the object with all contexts removed

defined_augment(*self*)
This fuction must return False if *identifier_augment()* would raise an IdentifierMissingException. Override it when defining a non-standard identifier for sub-classes of DataObjects.

get_owners(*self*, *property_class_name*)
Return a generator of owners along a property pointing to this object

graph_pattern(*self*, *shorten=False*, *show_namespaces=True*, ***kwargs*)
Get the graph pattern for this object.

It should be as simple as converting the result of triples() into a BGP

Parameters

shorten [bool] Indicates whether to shorten the URLs with the namespace manager attached to the self

id_is_variable(*self*)
Is the identifier a variable?

identifier_augment(*self*)
Override this method to define an identifier in lieu of one explicity set.

One must also override *defined_augment()* to return True whenever this method could return a valid identifier. IdentifierMissingException should be raised if an identifier cannot be generated by this method.

Raises

IdentifierMissingException

make_identifier_from_properties(*self*, *names*)
Creates an identifier from properties

retract(*self*)
Remove this object from the data store.

save(*self*)
Write in-memory data to the database. Derived classes should call this to update the store.

variable(*self*)
Must return a `Variable` object that identifies this `GraphObject` in queries.

The variable can be randomly generated when the object is created and stored in the object.

po_cache = None
A cache of property URIs and values. Used by RealSimpleProperty

properties_are_init_args = True
If true, then properties defined in the class body can be passed as keyword arguments to __init__. For example:

```
>>> class A(DataObject):
...     p = DatatypeProperty()

>>> A(p=5)
```

If the arguments are written explicitly into the __init__ method definition, then no special processing is done.

class owmeta.dataObject.**ContextMappedClass**(*name*, *bases*, *dct*)
Bases: yarom.mappedClass.MappedClass, *owmeta.contextualize.ContextualizableClass*

after_mapper_module_load(*self*, *mapper*)
Called after the module has been loaded. See *owmeta.mapper.Mapper*

contextualize_class_augment(*self*, *context*)
For MappedClass, rdf_type and rdf_namespace have special behavior where they can be auto-generated based on the class name and base_namespace. We have to pass through these values to our "proxy" to avoid this behavior

definition_context
Unlike self.context, definition_context isn't meant to be overriden

query
Stub. Eventually, creates a proxy that changes how some things behave for purposes of querying

class owmeta.dataObject.**DataObject**(***kwargs*)
Bases: *owmeta.dataObject.BaseDataObject*

owmeta.datasource module

exception owmeta.datasource.**DuplicateAlsoException**
Bases: Exception

class owmeta.datasource.**BaseDataTranslator**(***kwargs*)
Bases: *owmeta.dataObject.DataObject*

Translates from a data source to owmeta objects

input_type
alias of *DataSource*

output_type
alias of *DataSource*

translation_type
alias of *Translation*

make_translation(*self*, *sources=()*)
It's intended that implementations of DataTranslator will override this method to make custom Translations according with how different arguments to Translate are (or are not) distinguished.

The actual properties of a Translation subclass must be defined within the 'translate' method

translate(*self*, **args*, ***kwargs*)
Notionally, this method takes a data source, which is translated into some other data source. There doesn't necessarily need to be an input data source.

class owmeta.datasource.**DataObjectContextDataSource**(*context*, ***kwargs*)
Bases: *owmeta.datasource.DataSource*

Input source [ObjectProperty] Attribute: source

The data source that was translated into this one

Translation [ObjectProperty] Attribute: translation

Information about the translation process that created this object

Description [DatatypeProperty] Attribute: description

Free-text describing the data source

class owmeta.datasource.**DataSource**(***kwargs*)
Bases: *owmeta.dataObject.DataObject*

A source for data that can get translated into owmeta objects.

The value for any field can be passed to __init__ by name. Additionally, if the sub-class definition of a DataSource assigns a value for that field like:

```
class A(DataSource):
    some_field = 3
```

that value will be used over the default value for the field, but not over any value provided to __init__.

commit(*self*)
Commit the data source *locally*

This includes staging files such as they would be available for a translation. In general, a sub-class should implement commit_augment() rather than this method, or at least call this method via super

For example, if the data source produces a file, that file should be in

defined_augment(*self*)
This fuction must return False if *identifier_augment()* would raise an IdentifierMissingException. Override it when defining a non-standard identifier for sub-classes of DataObjects.

identifier_augment(*self*)
Override this method to define an identifier in lieu of one explicity set.

One must also override *defined_augment()* to return True whenever this method could return a valid identifier. IdentifierMissingException should be raised if an identifier cannot be generated by this method.

Raises

IdentifierMissingException

class owmeta.datasource.**DataSourceType**(*name*, *bases*, *dct*)
Bases: *owmeta.dataObject.ContextMappedClass*

A type for DataSources

Sets up the graph with things needed for MappedClasses

class owmeta.datasource.**DataTransatorType**(*name*, *bases*, *dct*)
Bases: *owmeta.dataObject.ContextMappedClass*

class owmeta.datasource.**DataTranslator**(***kwargs*)
Bases: *owmeta.datasource.BaseDataTranslator*

A specialization with the *GenericTranslation* translation type that adds sources for the translation automatically when a new output is made

translation_type
alias of *GenericTranslation*

make_translation(*self*, *sources=()*)
It's intended that implementations of DataTranslator will override this method to make custom Translations according with how different arguments to Translate are (or are not) distinguished.

The actual properties of a Translation subclass must be defined within the 'translate' method

class owmeta.datasource.**GenericTranslation**(***kwargs*)
Bases: *owmeta.datasource.Translation*

A generic translation that just has sources in order

defined_augment(*self*)
This fuction must return False if *identifier_augment()* would raise an `IdentifierMissingException`. Override it when defining a non-standard identifier for subclasses of DataObjects.

identifier_augment(*self*)
Override this method to define an identifier in lieu of one explicity set.

One must also override *defined_augment()* to return True whenever this method could return a valid identifier. `IdentifierMissingException` should be raised if an identifier cannot be generated by this method.

Raises

IdentifierMissingException

class owmeta.datasource.**OneOrMore**(*source_type*)
Bases: `object`

Wrapper for *DataTranslator* input *DataSource* types indicating that one or more of the wrapped type must be provided to the translator

class owmeta.datasource.**PersonDataTranslator**(***kwargs*)
Bases: *owmeta.datasource.BaseDataTranslator*

A person who was responsible for carrying out the translation of a data source manually

person
A person responsible for carrying out the translation.

class owmeta.datasource.**Translation**(***kwargs*)
Bases: *owmeta.dataObject.DataObject*

Representation of the method by which a DataSource was translated and the sources of that translation. Unlike the 'source' field attached to DataSources, the Translation may distinguish different kinds of input source to a translation.

defined_augment(*self*)
This fuction must return False if *identifier_augment()* would raise an `IdentifierMissingException`. Override it when defining a non-standard identifier for sub-classes of DataObjects.

identifier_augment(*self*)
Override this method to define an identifier in lieu of one explicity set.

One must also override *defined_augment()* to return True whenever this method could return a valid identifier. `IdentifierMissingException` should be raised if an identifier cannot be generated by this method.

Raises

IdentifierMissingException

owmeta.datasource_loader module

DataSourceLoaders take a data source identifier and retrieve the primary data (e.g., CSV files, electrode recordings) from some location (e.g., a file store, via a bittorrent tracker).

Each loader can treat the base_directory given as its own namespace and place directories in there however it wants.

exception owmeta.datasource_loader.**LoadFailed**(*data_source*, *loader*, **args*)
Bases: `Exception`

class owmeta.datasource_loader.**DataSourceDirLoader**(*base_directory=None*)
Bases: `object`

Loads a data files for a DataSource

The loader is expected to organize files for each data source within the given base directory.

can_load(*self*, *data_source*)
Returns true if the *DataSource* can be loaded by this loader

Parameters

data_source [*owmeta.datasource.DataSource*] The data source to load files for

load(*self*, *data_source*)
Loads the files for the data source

Parameters

data_source [*owmeta.datasource.DataSource*] The data source to load files for

Returns

A path to the loaded resource

owmeta.document module

exception owmeta.document.**PubmedRetrievalException**
Bases: `Exception`

exception owmeta.document.**WormbaseRetrievalException**
Bases: `Exception`

class owmeta.document.**BaseDocument**(***kwargs*)
Bases: *owmeta.dataObject.DataObject*

class owmeta.document.**Document**(*bibtex=None*, *doi=None*, *pubmed=None*, *wormbase=None*, ***kwargs*)
Bases: *owmeta.document.BaseDocument*

A representation of some document.

Possible keys include:

```
pmid, pubmed: a pubmed id or url (e.g., 24098140)
wbid, wormbase: a wormbase id or url (e.g., WBPaper00044287)
doi: a Digitial Object id or url (e.g., s00454-010-9273-0)
uri: a URI specific to the document, preferably usable for accessing
     the document
```

Parameters

bibtex [str] A string containing a single BibTeX entry. Parsed during initialization, but not saved thereafter. optional

doi [str] A Digital Object Identifier (DOI). optional

pubmed [str] A PubMed ID (PMID) or URL that points to a paper. Ignored if 'pmid' is provided. optional

wormbase [str] An ID or URL from WormBase that points to a record. Ignored if *wbid* or *wormbaseid* are provided. optional

defined_augment(*self*)
This fuction must return False if *identifier_augment()* would raise an IdentifierMissingException. Override it when defining a non-standard identifier for sub-classes of DataObjects.

identifier_augment(*self*)
Override this method to define an identifier in lieu of one explicity set.

One must also override *defined_augment()* to return True whenever this method could return a valid identifier. IdentifierMissingException should be raised if an identifier cannot be generated by this method.

Raises

IdentifierMissingException

update_from_wormbase(*self*, *replace_existing=False*)
Queries wormbase for additional data to fill in the Document.

If replace_existing is set to True, then existing values will be cleared.

author
An author of the document

date
Alias to year

doi
A Digital Object Identifier (DOI), optional

pmid
A PubMed ID (PMID) that points to a paper

title
The title of the document

uri
A non-standard URI for the document

wbid
An ID from WormBase.org that points to a record, optional

wormbaseid
An alias to *wbid*

year
The year (e.g., publication year) of the document

owmeta.documentContext module

class owmeta.documentContext.**DocumentContext**(*document*)
Bases: *owmeta.context.Context*

A Context that corresponds to a document.

class owmeta.documentContext.**DocumentContextMeta**
Bases: *owmeta.context.ContextMeta*

owmeta.evidence module

exception owmeta.evidence.**EvidenceError**
Bases: Exception

class owmeta.evidence.**Evidence**(***kwargs*)
Bases: *owmeta.dataObject.DataObject*

A representation which provides evidence, for a group of statements.

Attaching evidence to an set of statements is done like this:

```
>>> from owmeta.connection import Connection
>>> from owmeta.evidence import Evidence
>>> from owmeta.context import Context
```

Declare contexts:

```
>>> ACTX = Context(ident="http://example.org/data/some_statements")
>>> BCTX = Context(ident="http://example.org/data/some_other_statements")
>>> EVCTX = Context(ident="http://example.org/data/some_statements#evidence")
```

Make statements in ACTX and BCTX contexts:

```
>>> ACTX(Connection)(pre_cell="VA11", post_cell="VD12", number=3)
>>> BCTX(Connection)(pre_cell="VA11", post_cell="VD12", number=2)
```

In EVCTX, state that a that a certain document supports the set of statements in ACTX, but refutes the set of statements in BCTX:

```
>>> doc = EVCTX(Document)(author='White et al.', date='1986')
>>> EVCTX(Evidence)(reference=doc, supports=ACTX.rdf_object)
>>> EVCTX(Evidence)(reference=doc, refutes=BCTX.rdf_object)
```

Finally, save the contexts:

```
>>> ACTX.save_context()
>>> BCTX.save_context()
>>> EVCTX.save_context()
```

One note about the *reference* predicate: the reference should, ideally, be an unambiguous link to a peer-reviewed piece of scientific literature detailing methods and data analysis that supports the set of statements. However, in gather data from pre-existing sources, going to that level of specificity may be difficult due to deficient query capability at the data source. In such cases, a broader reference, such as a `Website` with information which guides readers to a peer-reviewed article supporting the statement is sufficient.

defined_augment(*self*)

This fuction must return False if *identifier_augment()* would raise an `IdentifierMissingException`. Override it when defining a non-standard identifier for subclasses of DataObjects.

identifier_augment(*self*)

Override this method to define an identifier in lieu of one explicity set.

One must also override *defined_augment()* to return True whenever this method could return a valid identifier. `IdentifierMissingException` should be raised if an identifier cannot be generated by this method.

Raises

IdentifierMissingException

reference

The resource providing evidence supporting/refuting the attached context

refutes

A context naming a set of statements which are refuted by the attached reference

supports

A context naming a set of statements which are supported by the attached reference

owmeta.evidence.**evidence_for**(*qctx*, *ctx*, *evctx=None*)

Returns an iterable of Evidence

Parameters

qctx [`object`] an object supported by evidence. If the object is a *Context* with no identifier, then the query considers statements 'staged' (rather than stored) in the context

ctx [`Context`] Context that bounds where we look for statements about `qctx`. The contexts for statements found in this context are the actual targets of Evidence.supports statements.

evctx [`Context`] if the Evidence.supports statements should be looked for somewhere other than `ctx`, that can be specified in evctx. optional

owmeta.evidence.**query_context**(*graph*, *qctx*)

graph [rdflib.graph.Graph] Graph where we can find the contexts for statements in `qctx`

qctx [owmeta.context.Context] Container for statements

owmeta.experiment module

class owmeta.experiment.**Experiment**(***kwargs*)
Bases: *owmeta.dataObject.DataObject*

Generic class for storing information about experiments

Should be overridden by specific types of experiments (example: see PatchClampExperiment in channelworm.py).

Overriding classes should have a list called "conditions" that contains the names of experimental conditions for that particular type of experiment. Each of the items in "conditions" should also be either a DatatypeProperty or ObjectProperty for the experiment as well.

get_conditions(*self*)
Return conditions and their associated values in a dict.

reference
Supporting article for this experiment.

owmeta.file_match module

owmeta.git_repo module

owmeta.identifier_mixin module

owmeta.identifier_mixin.**IdMixin**(*typ=<class 'object'>*, *hashfunc=None*)
Mixin that provides common identifier logic

Parameters

typ [type] The type of object to use as the hash function's super class. Defaults to 'object'

hashfunc [function] The function to use for encoding data provided to make_identifier. Should return an object can `.encode()` to a bytes (a.k.a. str in Python 2). Defaults to `hashlib.sha224()`

owmeta.import_contextualizer module

class owmeta.import_contextualizer.**ImportContextualizer**
Bases: object

Interface for classes that 'contextualize' an import.

Contextualizing an import means that if an object is defined with the name X in some context, and an import statement is written like this:

```
import X.a
```

X will be invoked like this:

```
a = X(__import__('a'))
```

On the other hand, for an import statement of this form:

```
from X.a import A
```

will cause X to be invoked as:

```
_temp = X(__import__('a', globals(), locals(), ('A',)), ('A',))
A = _temp.A
```

and the import statement:

```
from X.a import A as AA
```

will cause X to be invoked as:

```
_temp = X(__import__('a', globals(), locals(), ('A',)), ('A',))
AA = _temp.A
```

meaning that the contextualizer won't know what the 'as' name is.

For the astute reader, you may notice parallels between the protocol for ImportContextualizer and the __import__ function itself. You may also be wondering why the contextualizer isn't merely a proxy for __import__. The first reason is that I want the true import to always happen, regardless of what the contextualizer does, so that the semantics of the contextualizer are very clear. The second reason is that requiring a real proxy would require more complexity in the contextualizers to ensure that they are properly handling exceptions, module attributes and the return value of __import__, ensuring that the *right* __import__ is used, as well as handling the 'fromlist' correctly.

owmeta.import_override module

owmeta.inverse_property module

For declaring inverse properties of GraphObjects

exception owmeta.inverse_property.**InversePropertyException**
Bases: Exception

class owmeta.inverse_property.**InversePropertyMixin**
Bases: object

Mixin for inverse properties.

Augments RealSimpleProperty methods to update inverse properties as well

owmeta.mapper module

exception owmeta.mapper.**UnmappedClassException**
Bases: Exception

class owmeta.mapper.**Mapper**(*base_namespace=None*, *imported=()*, *name=None*, ***kwargs*)
Bases: owmeta.module_recorder.ModuleRecordListener, *owmeta.configure.Configureable*

Keeps track of relationships between classes, between modules, and between classes and modules

decorate_class(*self*, *cls*)
Extension point for subclasses of Mapper to apply an operation to all mapped classes

load_module(*self*, *module_name*)
Loads the module.

lookup_class(*self*, *cname*)
Gets the class corresponding to a fully-qualified class name

DecoratedMappedClasses = None
Maps RDF types to properties of the related class

MappedClasses = None
Maps classes to decorated versions of the class

base_namespace = None
Modules that have already been loaded

owmeta.module_recorder module

owmeta.muscle module

class owmeta.muscle.**Muscle**(*name=None*, *lineageName=None*, ***kwargs*)
Bases: *owmeta.cell.Cell*

A single muscle cell.

See what neurons innervate a muscle:

Example:

```
>>> mdr21 = Muscle('MDR21')
>>> innervates_mdr21 = mdr21.innervatedBy()
>>> len(innervates_mdr21)
4
```

innervatedBy
Neurons synapsing with this muscle

neurons
Alias to *innervatedBy*

receptor
Alias to *receptors*

receptors
Receptor types expressed by this type of muscle

owmeta.my_neuroml module

class owmeta.my_neuroml.**NeuroML**(**args*, ***kwargs*)
Bases: *owmeta.data.DataUser*

classmethod generate(*o*, *t=2*)
Get a NeuroML object that represents the given object. The `type` determines what content is included in the NeuroML object:

Parameters

- **o** – The object to generate neuroml from
- **t** – The what kind of content should be included in the document - 0=full morphology+biophysics - 1=cell body only+biophysics - 2=full morphology only

Returns A NeuroML object that represents the given object.

Return type NeuroMLDocument

classmethod write(*o*, *n*)
Write the given neuroml document object out to a file :param o: The NeuroMLDocument to write :param n: The name of the file to write to

owmeta.network module

class owmeta.network.**Network**(*worm=None*, ***kwargs*)
Bases: *owmeta.biology.BiologyType*

A network of neurons

aneuron(*self*, *name*)
Get a neuron by name.

Example:

```
# Grabs the representation of the neuronal network
>>> net = Worm().get_neuron_network()

# Grab a specific neuron
>>> aval = net.aneuron('AVAL')

>>> aval.type()
set([u'interneuron'])
```

Parameters **name** – Name of a c. elegans neuron

Returns Neuron corresponding to the name given

Return type *owmeta.neuron.Neuron*

defined_augment(*self*)
This fuction must return False if *identifier_augment()* would raise an `IdentifierMissingException`. Override it when defining a non-standard identifier for subclasses of DataObjects.

identifier_augment(*self*)
Override this method to define an identifier in lieu of one explicity set.

One must also override *defined_augment()* to return True whenever this method could return a valid identifier. `IdentifierMissingException` should be raised if an identifier cannot be generated by this method.

Raises

IdentifierMissingException

interneurons(*self*)
Get all interneurons

Returns A iterable of all interneurons

Return type iter(*Neuron*)

motor(*self*)
Get all motor

Returns A iterable of all motor neurons

Return type iter(*Neuron*)

neuron_names(*self*)
Gets the complete set of neurons' names in this network.

Example:

```
# Grabs the representation of the neuronal network
>>> net = Worm().get_neuron_network()

#NOTE: This is a VERY slow operation right now
>>> len(set(net.neuron_names()))
302
>>> set(net.neuron_names())
set(['VB4', 'PDEL', 'HSNL', 'SIBDR', ... 'RIAL', 'MCR', 'LUAL'])
```

sensory(*self*)
Get all sensory neurons

Returns A iterable of all sensory neurons

Return type iter(*Neuron*)

neuron
Returns a set of all Neuron objects in the network

neurons
Alias to *neuron*

synapse
Returns a set of all synapses in the network

synapses
Alias to *synapse*

worm
The worm connected to the network

owmeta.neuron module

class owmeta.neuron.**ConnectionProperty**(***kwargs*)
Bases: *owmeta.pProperty.Property*

A representation of the connection between neurons. Either a gap junction or a chemical synapse

TODO: Add neurotransmitter type. TODO: Add connection strength

This is defined so that the __init__ method gets a contextualized instance, allowing for statements made in __init__ to be contextualized.

get(*self*, *pre_post_or_either='pre'*, ***kwargs*)
Get a list of connections associated with the owning neuron.

Parameters

pre_post_or_either: str What kind of connection to look for. 'pre': Owner is the source of the connection 'post': Owner is the destination of the connection 'either': Owner is either the source or destination of the connection

Returns

list of Connection

set(*self*, *conn*, ***kwargs*)
Add a connection associated with the owner Neuron

Parameters

conn [*owmeta.connection.Connection*] connection associated with the owner neuron

Returns

A owmeta.neuron.Connection

class owmeta.neuron.**Neighbor**(***kwargs*)

Bases: *owmeta.pProperty.Property*

This is defined so that the __init__ method gets a contextualized instance, allowing for statements made in __init__ to be contextualized.

get(*self*, ***kwargs*)

Get a list of neighboring neurons.

Parameters

See parameters for owmeta.connection.Connection

Returns

list of *Neuron*

set(*self*, *other*, ***kwargs*)

Set the value of this property

Derived classes must override.

class owmeta.neuron.**Neuron**(*name=False*, ***kwargs*)

Bases: *owmeta.cell.Cell*

A neuron.

See what neurons express some neuropeptide

Example:

```
# Grabs the representation of the neuronal network
>>> net = P.Worm().get_neuron_network()

# Grab a specific neuron
>>> aval = net.aneuron('AVAL')

>>> aval.type()
set([u'interneuron'])

#show how many connections go out of AVAL
>>> aval.connection.count('pre')
77

>>> aval.name()
u'AVAL'

#list all known receptors
>>> sorted(aval.receptors())
[u'GGR-3', u'GLR-1', u'GLR-2', u'GLR-4', u'GLR-5', u'NMR-1', u'NMR-2', u'UNC-8']

#show how many chemical synapses go in and out of AVAL
>>> aval.Syn_degree()
90
```

Parameters

name [str] The name of the neuron.

Attributes

neighbor [Property] Get neurons connected to this neuron if called with no arguments, or with arguments, state that neuronName is a neighbor of this Neuron

connection [Property] Get a set of Connection objects describing chemical synapses or gap junctions between this neuron and others

GJ_degree(*self*)
Get the degree of this neuron for gap junction edges only

Returns total number of incoming and outgoing gap junctions

Return type int

Syn_degree(*self*)
Get the degree of this neuron for chemical synapse edges only

Returns total number of incoming and outgoing chemical synapses

Return type int

get_incidents(*self*, *type=0*)
Get neurons which synapse at this neuron

innexin
Innexin types associated with this neuron

neuropeptide
Name of the gene corresponding to the neuropeptide produced by this neuron

neurotransmitter
Neurotransmitters associated with this neuron

receptor
The receptor types associated with this neuron

receptors
Alias to py:attr:*receptor*

type
The neuron type (i.e., sensory, interneuron, motor)

owmeta.pProperty module

class owmeta.pProperty.**Property**(*name=False*, *owner=False*, ***kwargs*)
Bases: *owmeta.contextualize.Contextualizable*, *owmeta.data.DataUser*

Store a value associated with a DataObject

Properties can be be accessed like methods. A method call like:

```
a.P()
```

for a property P will return values appropriate to that property for a, the owner of the property.

Parameters

owner [*owmeta.dataObject.DataObject*] The owner of this property

name [str] The name of this property. Can be accessed as an attribute like:

```
owner.name
```

This is defined so that the __init__ method gets a contextualized instance, allowing for statements made in __init__ to be contextualized.

get (*self*, **args*)
Get the things which are on the other side of this property

The return value must be iterable. For a `get` that just returns a single value, an easy way to make an iterable is to wrap the value in a tuple like `(value,)`.

Derived classes must override.

has_value (*self*)
Returns true if the Property has any values set on it.

This may be defined differently for each property

one (*self*)
Returns a single value for the `Property` whether or not it is multivalued.

set (*self*, **args*, ***kwargs*)
Set the value of this property

Derived classes must override.

class owmeta.pProperty.**PropertyMeta** (*name*, *bases*, *dct*)
Bases: *owmeta.contextualize.ContextualizableClass*

owmeta.package_utils module

owmeta.plot module

class owmeta.plot.**Plot** (*data=None*, **args*, ***kwargs*)
Bases: *owmeta.dataObject.DataObject*

Object for storing plot data in owmeta.

Parameters

data [2D list (list of lists)] List of XY coordinates for this Plot.

Example usage ::

```
>>> pl = Plot([[1, 2], [3, 4]])
>>> pl.get_data()
# [[1, 2], [3, 4]]
```

get_data (*self*)
Get the data stored for this plot.

set_data (*self*, *data*)
Set the data attribute, which is user-facing, as well as the serialized _data_string attribute, which is used for db storage.

owmeta.rdf_go_modifiers module

owmeta.rdf_query_util module

owmeta.rdf_query_util.**get_most_specific_rdf_type**(*types*, *context=None*, *bases=()*)
Gets the most specific rdf_type.

Returns the URI corresponding to the lowest in the DataObject class hierarchy from among the given URIs.

owmeta.rdf_query_util.**oid**(*identifier_or_rdf_type=None*, *rdf_type=None*, *context=None*, *base_type=None*)
Create an object from its rdf type

Parameters

identifier_or_rdf_type [str or rdflib.term.URIRef] If rdf_type is provided, then this value is used as the identifier for the newly created object. Otherwise, this value will be the rdf_type of the object used to determine the Python type and the object's identifier will be randomly generated.

rdf_type [str, rdflib.term.URIRef, False] If provided, this will be the rdf_type of the newly created object.

Returns

The newly created object

owmeta.simpleProperty module

class owmeta.simpleProperty.**ContextMappedPropertyClass**(**args*, ***kwargs*)
Bases: yarom.mappedProperty.MappedPropertyClass, *owmeta.contextualize.ContextualizableClass*

class owmeta.simpleProperty.**ContextualizedPropertyValue**(*value*)
Bases: yarom.propertyValue.PropertyValue

class owmeta.simpleProperty.**DatatypeProperty**(*resolver*, ***kwargs*)
Bases: yarom.propertyMixins.DatatypePropertyMixin, owmeta.simpleProperty.PropertyCountMixin, *owmeta.simpleProperty.RealSimpleProperty*

Parameters

resolver [RDFTypeResolver] Resolves RDF identifiers returned from *get()* into objects

get(*self*)
Gets a config value from this Configureable's conf

See also:

Configure.get

class owmeta.simpleProperty.**ObjectProperty**(*resolver=None*, **args*, ***kwargs*)
Bases: *owmeta.inverse_property.InversePropertyMixin*, owmeta.simpleProperty._ContextualizingPropertySetMixin, owmeta.simpleProperty.PropertyCountMixin, *owmeta.simpleProperty.RealSimpleProperty*

This is defined so that the __init__ method gets a contextualized instance, allowing for statements made in __init__ to be contextualized.

get(*self*)
Gets a config value from this `Configureable`'s `conf`

See also:

Configure.get

class owmeta.simpleProperty.**POCache**
Bases: `tuple`

The predicate-object cache object

class owmeta.simpleProperty.**RealSimpleProperty**(*owner*, ***kwargs*)
Bases: *owmeta.data.DataUser*, *owmeta.contextualize.Contextualizable*

This is defined so that the __init__ method gets a contextualized instance, allowing for statements made in __init__ to be contextualized.

clear(*self*)
Clears values set *in all contexts*

decontextualize(*self*)
Return the object with all contexts removed

get(*self*)
Gets a config value from this `Configureable`'s `conf`

See also:

Configure.get

class owmeta.simpleProperty.**UnionProperty**(*resolver*, ***kwargs*)
Bases: `owmeta.simpleProperty._ContextualizingPropertySetMixin`, *owmeta.inverse_property.InversePropertyMixin*, `yarom.propertyMixins.UnionPropertyMixin`, `owmeta.simpleProperty.PropertyCountMixin`, *owmeta.simpleProperty.RealSimpleProperty*

A Property that can handle either DataObjects or basic types

Parameters

resolver [`RDFTypeResolver`] Resolves RDF identifiers into objects returned from *get()*

get(*self*)
Gets a config value from this `Configureable`'s `conf`

See also:

Configure.get

owmeta.statement module

class owmeta.statement.**Statement**(*subject*, *property*, *object*, *context*)
Bases: *owmeta.statement.Statement*

Create new instance of Statement(subject, property, object, context)

owmeta.text_util module

owmeta.utils module

Common utilities for translation, massaging data, etc., that don't fit elsewhere in owmeta

owmeta.utils.**grouper**(*iterable*, *n*, *fillvalue=None*)
Collect data into fixed-length chunks or blocks

owmeta.website module

class owmeta.website.**Website**(*title=None*, ***kwargs*)
Bases: *owmeta.document.BaseDocument*

A representation of a website

defined_augment(*self*)
This fuction must return False if *identifier_augment()* would raise an IdentifierMissingException. Override it when defining a non-standard identifier for sub-classes of DataObjects.

identifier_augment(*self*)
Override this method to define an identifier in lieu of one explicity set.

One must also override *defined_augment()* to return True whenever this method could return a valid identifier. IdentifierMissingException should be raised if an identifier cannot be generated by this method.

Raises

IdentifierMissingException

title
The official name for the website

url
A URL for the website

owmeta.worm module

class owmeta.worm.**Worm**(*scientific_name=False*, ***kwargs*)
Bases: *owmeta.biology.BiologyType*

A representation of the whole worm

defined_augment(*self*)
True if the name is defined

get_neuron_network(*self*)
Return the neuron network of the worm.

Example:

```
# Grabs the representation of the neuronal network
>>> net = P.Worm().get_neuron_network()

# Grab a specific neuron
>>> aval = net.aneuron('AVAL')
```

(continues on next page)

(continued from previous page)

```
>>> aval.type()
set([u'interneuron'])

#show how many connections go out of AVAL
>>> aval.connection.count('pre')
77
```

Returns An object to work with the network of the worm

Return type owmeta.Network

get_semantic_net(*self*)

Get the underlying semantic network as an RDFLib Graph

Returns A semantic network containing information about the worm

Return type rdflib.ConjunctiveGraph

identifier_augment(*self*, **args*, ***kwargs*)

Result is derived from the name property

muscles(*self*)

Get all Muscle objects attached to the Worm.

Example:

```
>>> muscles = P.Worm().muscles()
>>> len(muscles)
96
```

Returns A set of all muscles

Return type set

cell

A cell in the worm

muscle

A type of muscle which is in the worm

name

Alias to *scientific_name*

neuron_network

The neuron network of the worm

scientific_name

Scientific name for the organism

owmeta.worm_common module

CHAPTER 2

For Users

2.1 owmeta Data Sources

The sources of data for owmeta are stored in the OpenWormData repository. A few `DataTranslators` translate these data into common owmeta data sources. You can list these by running:

```
owm source list
```

and you can show some of the properties of a data source by running:

```
owm source show $SOURCE_IDENTIFIER
```

For instance, you can run the following to see the top-level data source, try:

```
owm source show http://openworm.org/data
```

This will print out summary descriptions of the sources that contribute to the main data source.

2.1.1 A Note on owmeta Data

Below, each major element of the worm's anatomy that owmeta stores data on is considered individually. The data being used is tagged by source in a superscript, and the decisions made during the curation process (if any) are described.

2.1.2 Neurons

- Neuron names[2]: Extracted from WormBase. Dynamic version on this google spreadsheet. Staged in this csv file. Parsed by this method.

[2]
- Harris, T. W., Antoshechkin, I., Bieri, T., Blasiar, D., Chan, J., Chen, W. J., … Sternberg, P. W. (2010). WormBase: a comprehensive resource for nematode research. Nucleic Acids Research, 38(Database issue), D463–7. http://doi.org/10.1093/nar/gkp952
- Lee, R. Y. N., & Sternberg, P. W. (2003). Building a cell and anatomy ontology of Caenorhabditis elegans. Comparative and Functional

- Neuron types[1]: Extracted from WormAtlas.org. Staged in this csv file. Parsed by this method.
- Cell descriptions[1]: Extracted from WormAtlas.org. Staged in this tsv file. Parsed by this method.
- Lineage names[1]: Extracted from WormAtlas.org. Dynamic version on this google spreadsheet. Staged in this tsv file. Parsed by this method.
- Neurotransmitters[1]: Extracted from WormAtlas.org. Dynamic version on this google spreadsheet. Staged in this csv file. Parsed by this method.
- Neuropeptides[1]: Extracted from WormAtlas.org. Dynamic version on this google spreadsheet. Staged in this csv file. Parsed by this method.
- Receptors[1]: Extracted from WormAtlas.org. Dynamic version on this google spreadsheet. Staged in this csv file. Parsed by this method.
- Innexins[1]: Extracted from WormAtlas.org. Dynamic version on this google spreadsheet. Staged in this csv file. Parsed by this method.

Gene expression data below, additional to that extracted from WormAtlas concerning receptors, neuropeptides, neurotransmitters and innexins are parsed by this method:

- Monoamine secretors and receptors, neuropeptide secretors and receptors[4]: Dynamic version on this google spreadsheet. Staged in this csv file.

2.1.3 Muscle cells

- Muscle names[2]: Extracted from WormBase. Dynamic version on this google spreadsheet. Staged in this csv file. Parsed by this method.
- Cell descriptions[1]: Extracted from WormAtlas.org. Dynamic version on this google spreadsheet. Staged in this tsv file. Parsed by this method.
- Lineage names[1]: Extracted from WormAtlas.org. Dynamic version on this google spreadsheet. Staged in this tsv file. Parsed by this method.
- Neurons that innervate each muscle[3]: Extracted from data personally communicated by S. Cook. Staged in this csv file. Parsed by this method.

2.1.4 Connectome

- Gap junctions between neurons[3]: Extracted from data personally communicated by S. Cook. Staged in this csv file. Parsed by this method.
- Synapses between neurons[3]: Extracted from data personally communicated by S. Cook. Staged in this csv file. Parsed by this method.

Genomics, 4(1), 121–6. http://doi.org/10.1002/cfg.248

[1] Altun, Z.F., Herndon, L.A., Wolkow, C.A., Crocker, C., Lints, R. and Hall, D. H. (2015). WormAtlas. Retrieved from http://www.wormatlas.org - WormAtlas Complete Cell List

[4] Bentley B., Branicky R., Barnes C. L., Chew Y. L., Yemini E., Bullmore E. T., Vertes P. E., Schafer W. R. (2016) The Multilayer Connectome of Caenorhabditis elegans. PLoS Comput Biol 12(12): e1005283. http://doi.org/10.1371/journal.pcbi.1005283

[3] Emmons, S., Cook, S., Jarrell, T., Wang, Y., Yakolev, M., Nguyen, K., Hall, D. Whole-animal C. elegans connectomes. C. Elegans Meeting 2015 http://abstracts.genetics-gsa.org/cgi-bin/celegans15s/wsrch15.pl?author=emmons&sort=ptimes&sbutton=Detail&absno=155110844&sid=668862

Curation note

There was another source of *C. elegans* connectome data that was created by members of the OpenWorm project that has since been retired. The history of this spreadsheet is mostly contained in this forum post We decided to use the Emmons data set[3] as the authoritative source for connectome data, as it is the very latest version and updated version of the *C. elegans* connectome that we are familiar with.

2.1.5 Data Source References

2.2 Requirements for data storage in OpenWorm

Our OpenWorm database captures facts about `C. elegans`. The database stores data for generating model files and together with annotations describing the origins of the data. Below are a set of recommendations for implementation of the database organized around an RDF model.

2.2.1 Interface

Access is through a Python library which communicates with the database. This library serves the function of providing an object oriented view on the database that can be accessed through the Python scripts commonly used in the project. The *api* is described separately.

2.2.2 Data modelling

Biophysical and anatomical data are included in the database. A sketch of some features of the data model is below. Also included in our model are the relationships between these types. Given our choice of data types, we do not model the individual interactions between cells as entities in the database. Rather these are described by generic predicates in an RDF triple. For instance, neuron A synapsing with muscle cell B would give a statement (A, synapsesWith, B), but A synapsing with neuron C would also have (A, synapsesWith, C). Data which belong to the specific relationship between two nodes is attached to an rdf:Statement object which points to the statement. This choice is intended to easy querying and extension later on.

Nervous system

In the worm's nervous system, we capture a few important data types (listed *below*). These correspond primarily to the anatomical structures and chemicals which are necessary for the worm to record external and internal stimuli and activate its body in response to those stimuli.

Data types

A non-exhaustive list of neurological data types in our C. elegans database:

- receptor types identified in the nerve cell
- neurons
- ion channels
- neurotransmitters
- muscle receptors

Development

Caenorhabditis elegans has very stable cell division patterns in the absence of mutations. This means that we can capture divisions in our database as static 'daughter_of' relationships. The theory of differentiation codes additionally gives an algorithmic description to the growth patterns of the worm which describes signals transmitted between developing cells. In order to test this theory we would like to leverage existing photographic data indicating the volume of cells at the time of their division as this relates to the differentiation code stored by the cell. Progress on this issue is documented on Github.

Aging

Concurrently with development, we would like to begin modeling the effects of aging on the worm. Aging typically manifests in physiological changes due to transcription errors or cell death. These physiological changes can be represented abstractly as parameters to the function of biological entities. See Github for further discussion.

2.2.3 Information assurance

Reasoning and Data integrity

To make full use of RDF storage it's recommended to leverage reasoning over our stored data. Encoding rules for the worm requires a good knowledge of both C. elegans and the database schema. More research needs to be done on this going forward. Preliminarily, SPIN, a constraint notation system based on SPARQL looks like a good candidate for `specifying` rules, but an inference engine for `enforcing` the rules still needs to be found.

Input validation

Input validation is to be handled through the interface library referenced *above*. In general, incorrect entry of biological names will result in an error being reported identifying the offending entry and providing a acceptable entries where appropriate. No direct access to the underlying data store will be provided.

Provenance

Tracking the origins of facts stated in the database demands a method of annotating statements in our database. Providing citations for facts must be as simple as providing a global identifier (e.g., URI, DOI) or a local identifier (e.g., Bibtex identifier, Pubmed ID). A technique called RDF reification allows us to annotate arbitrary facts in our database with additional information. This technique allows for the addition of structured citation data to facts in the database as well as annotations for tracking responsibility for uploads to the database. Further details for the attachment of evidence using this technique are given in the *api*.

In line with current practices for communication through the source code management platform, Github, we would like to track responsibility for new uploads to the database. Two methods are proposed for tracking this information: RDF named graphs and RDF reification. Tracking information must include, at least, a time-stamp on the update and linking of the submitted data to the uploader's unique identifier (e.g., email address). Named graphs have the advantage of wide support for the use of tracking uploads. The choice between these depends largely the support of the chosen data store for named graphs.

Access control

Write access to data in the project has been inconsistent between various data sources in the project. Going forward, write access to OpenWorm databases should be restricted to authenticated users to forestall the possibility of malicious tampering.

One way to accomplish this would be to leverage GitHub's fork and pull model with the data as well as the code. This would require two things:

- Instead of remote hosting of data, data is local to each copy of the library within a local database
- A serialization method dumps a new copy of the data out to a flat file enabling all users of the library to contribute their modifications to the data back to the owmeta project via GitHub.

A follow on to #2 is that the serialization method would need to preserve the ordering of data elements and write in some plain text format so that a simple diff on GitHub would be able to illuminate changes that were made.

2.2.4 Miscellaneous

Versioning

Experimental methods are constantly improving in biological research. These improvements may require updating the data we reference or store internally. However, in making updates we must not immediately expunge older content, breaking links created by internal and external agents. Ideally we would have a means of deprecating old data and specifying replacements. On the level of single resources, this is a trivial mapping which may be done transparently to all readers. For a more significant change, altering the schema, human intervention may be required to update external readers.

2.2.5 Why RDF?

RDF offers advantages in resilience to schema additions and increased flexibility in integrating data from disparate sources.[1] These qualities can be valued by comparison to relational database systems. Typically, schema changes in a relational database require extensive work for applications using it.[2] In the author's experience, RDF databases offer more freedom in restructuring. Also, for data integration, SPARQL, the standard language for querying over RDF has Federated queries which allow for nearly painless integration of external SPARQL endpoints with existing queries.

FuXi

FuXi is implemented as a semantic reasoning layer in owmeta. In other words, it will be used to automatically infer (and set) properties from other properties in the worm database. This means that redundant information (ex: explicitly stating that each object is of class "dataType") and subclass relationships (ex: that every object of type "Neuron" is also of type "Cell"), as well as other relationships, can be generated by the firing of FuXi's rule engine, without being hand-coded.

Aside from the time it saves in coding, FuXi may allow for a smaller footprint in the cloud, as many relationships within the database could be inferred *after* download.

A rule might be:

- { x is "Neuron" } => { x is "Cell" }

And a fact might be:

- { "ADLR" is "Neuron" }

Given the above rule and fact, FuXi could infer the new fact:

- { "ADLR" is "Cell" }

[1] http://answers.semanticweb.com/questions/19183/advantages-of-rdf-over-relational-databases
[2] http://research.microsoft.com/pubs/118211/andy%20maule%20-%20thesis.pdf

The advantage of local storage of the database that goes along with each copy of the library is that the data will have the version number of the library. This means that data can be 'deprecated' along with a deprecated version of the library. This also will prevent changes made to a volatile database that break downstream code that uses the library.

2.3 Adding Data to *YOUR* OpenWorm Database

So, you've got some biological data about the worm and you'd like to save it in owmeta, but you don't know how it's done?

You've come to the right place!

A few biological entities (e.g., Cell, Neuron, Muscle, Worm) are pre-coded into owmeta. The full list is available in the *API*. If these entities already cover your use-case, then all you need to do is add values for the appropriate fields and save them. If you have data already loaded into your database, then you can load objects from it:

```
>>> from owmeta.neuron import Neuron
>>> n = Neuron.query()
>>> n.receptor('UNC-13')
owmeta.statement.Statement(...obj=yarom.propertyValue.PropertyValue(rdflib.term.
↪Literal(u'UNC-13')), context=None)
>>> for x in n.load():
...     do_something_with_unc13_neuron(n)  # doctest.SKIP
```

If you need additional entities it's easy to create them. Documentation for this is provided *here*.

Typically, you'll want to attach the data that you insert to entities already in the database. This allows you to recover objects in a hierarchical fashion from the database later. *Worm*, for instance, has a property, `neuron_network`, which points to the *Network* which should contain all neural cells and synaptic connections. To initialize the hierarchy you would do something like:

```
>>> from owmeta.context import Context
>>> from owmeta.worm import Worm
>>> from owmeta.network import Network
>>> ctx = Context('http://example.org/c-briggsae')
>>> w = ctx(Worm)('C. briggsae') # The name is optional and currently defaults to 'C.␣
↪elegans'
>>> nn = ctx(Network)()           # make a neuron network
>>> w.neuron_network(nn)          # attach to the worm the neuron network
owmeta.statement.Statement(...)
>>> n = ctx(Neuron)('NeuronX')    # make a neuron
>>> n.receptor('UNC-13')          # state that the neuron has a UNC-13 type receptor
owmeta.statement.Statement(...)
>>> nn.neuron(n)                  # attach to the neuron network
owmeta.statement.Statement(...)
>>> ctx.save()           # save all of the data attached to the worm
```

It is possible to create objects without attaching them to anything and they can still be referenced by calling load on an instance of the object's class as in `n.load()` above. This also points out another fact: you don't have to set up the hierarchy for each insert in order for the objects to be linked to existing entities. If you have previously set up connections to an entity (e.g., `Worm('C. briggsae')`), assuming you *only* have one such entity, you can refer to things attached to it without respecifying the hierarchy for each script. The database packaged with owmeta should have only one Worm and one Network.

Remember that once you've made all of the statements, you must save the context in which the statements are made.

Future capabilities:

- Adding propositional logic to support making statements about all entities matching some conditions without needing to `load()` and `save()` them from the database.
- Statements like:

```
ctx = Context('http://example.org/c-briggsae')
w = ctx.stored(Worm)()
w.neuron_network.neuron.receptor('UNC-13')
l = list(w.load()) # Get a list of worms with neurons expressing 'UNC-13'
```

  currently, to do the equivalent, you must work backwards, finding all neurons with UNC-13 receptors, then getting all networks with those neurons, then getting all worms with those networks:

```
worms = set()
n = ctx.stored(Neuron)()
n.receptor('UNC-13')
for ns in n.load():
    nn = ctx.stored(Network)()
    nn.neuron(ns)
    for z in nn.load():
        w = ctx.stored(Worm)()
        w.neuron_network(z)
        worms.add(w)
l = list(worms)
```

  It's not difficult logic, but it's 8 extra lines of code for a, conceptually, very simple query.

- Also, queries like:

```
l = list(ctx.stored(Worm)('C. briggsae').neuron_network.neuron.receptor()) # get␣
↪a list
#of all receptors expressed in neurons of C. briggsae
```

  Again, not difficult to write out, but in this case it actually gives a much longer query time because additional values are queried in a `load()` call that are never returned.

  We'd also like operators for composing many such strings so:

```
ctx.stored(Worm)('C. briggsae').neuron_network.neuron.get('receptor', 'innexin')
↪# list
#of (receptor, innexin) values for each neuron
```

  would be possible with one query and thus not requiring parsing and iterating over neurons twice–it's all done in a single, simple query.

2.3.1 Contexts

Above, we used contexts without explaining them. In natural languages, our statements are made in a context that influences how they should be interpreted. In owmeta, that kind of context-sensitivity is modeled by using *owmeta.context.Context* objects. To see what this looks like, let's start with an example.

Basics

I have data about widgets from BigDataWarehouse (BDW) that I want to translate into RDF using owmeta, but I don't want put them with my other widget data since BDW data may conflict with mine. Also, if get more BDW data, I want to be able to relate these data to that. A good way to keep data which are made at distinct times or which come from different, possibly conflicting, sources is using contexts. The code below shows how to do that:

```
>>> from rdflib import ConjunctiveGraph
>>> from owmeta.context import Context
>>> # from mymod import Widget  # my own OWM widget model
>>> # from bdw import Load # BigDataWarehouse API

>>> # Create a Context with an identifier appropriate to this BDW data import
>>> ctx = Context('http://example.org/data/imports/BDW_Widgets_2017-2018')

>>> # Create a context manager using the default behavior of reading the
>>> # dictionary of current local variables
>>> with ctx(W=Widget) as c:
...     for record in Load(data_set='Widgets2017-2018'):
...         # declares Widgets in this context
...         c.W(part_number=record.pnum,
...             fullness=record.flns,
...             hardiness=record.hrds)
Widget(ident=rdflib.term.URIRef(...))


>>> # Create an RDFLib graph as the target for the data
>>> g = ConjunctiveGraph()

>>> # Save the data
>>> ctx.save(g)

>>> # Serialize the data in the nquads format so we can see that all of our
>>> # statements are in the proper context
>>> print(g.serialize(format='nquads').decode('UTF-8'))
<http://openworm.org/entities/Widget/12> <http...> <http://example.org/data/imports/
↪BDW_Widgets_2017-2018> .
<http://openworm.org/entities/Widget/12> <...
```

If you've worked with lots of data before, this kind of pattern should be familiar. You can see how, with later imports, you would follow the naming scheme to create new contexts (e.g., `http://example.org/data/imports/BDW_Widgets_2018-2019`). These additional contexts could then have separate metadata attached to them or they could be compared:

```
>>> len(list(ctx(Widget)().load()))
1
>>> len(list(ctx18(Widget)().load()))  # 2018-2019 context
3
```

Context Metadata

Contexts, because they have identifiers just like any other objects, so we can make statements about them as well. An essential statement is imports: Contexts import other contexts, which means, if you follow owmeta semantics, that when you query objects from the importing context, that the imported contexts will also be available to query.

2.4 Making data objects

To make a new object type like *Neuron* or *owmeta.worm.Worm*, for the most part, you just need to make a Python class.

Say, for example, that I want to record some information about drug reactions in C. elegans. I make `Drug` and `Experiment` classes to describe C. elegans reactions:

```
>>> from owmeta.dataObject import (DataObject,
...                                DatatypeProperty,
...                                ObjectProperty,
...                                Alias)
>>> from owmeta.worm import Worm
>>> from owmeta.evidence import Evidence
>>> from owmeta.document import Document
>>> from owmeta.context import Context
>>> from owmeta.mapper import Mapper
>>> from owmeta import connect, ModuleRecorder

>>> class Drug(DataObject):
...     name = DatatypeProperty()
...     drug_name = Alias(name)
...     def identifier_augment(self):
...         return self.make_identifier_direct(self.name.onedef())
...
...     def defined_augment(self):
...         return self.name.has_defined_value()

>>> class Experiment(DataObject):
...     drug = ObjectProperty(value_type=Drug)
...     subject = ObjectProperty(value_type=Worm)
...     route_of_entry = DatatypeProperty()
...     reaction = DatatypeProperty()

# Do some accounting stuff to register the classes. Usually happens behind
# the scenes.
>>> m = Mapper()
>>> ModuleRecorder.add_listener(m)
>>> m.process_classes(Drug, Experiment)
```

So, we have created I can then make a Drug object for moon rocks and describe an experiment by Aperture Labs:

```
>>> ctx = Context('http://example.org/experiments', mapper=m)
>>> d = ctx(Drug)(name='moon rocks')
>>> e = ctx(Experiment)(key='experiment001')
>>> w = ctx(Worm)('C. elegans')
>>> e.subject(w)
owmeta.statement.Statement(...Context(.../experiments"))

>>> e.drug(d)
owmeta.statement.Statement(...)

>>> e.route_of_entry('ingestion')
owmeta.statement.Statement(...)

>>> e.reaction('no reaction')
owmeta.statement.Statement(...)

>>> ev = Evidence(key='labresults', reference=Document(author="Aperture Labs"))
>>> ev.supports(ctx)
owmeta.statement.Statement(...)
```

and save those statements:

```
>>> ctx.save()
```

For simple objects, this is all we have to do.

You can also add properties to an object after it has been created by calling either ObjectProperty or DatatypeProperty on the class:

```
>>> d = ctx(Drug)(name='moon rocks')
>>> Drug.DatatypeProperty('granularity', owner=d)
__main__.Drug_granularity(owner=Drug(ident=rdflib.term.URIRef(u'http://openworm.org/
↪entities/Drug/moon%20rocks')))

>>> d.granularity('ground up')
owmeta.statement.Statement(...Context(.../experiments"))

>>> do = Drug()
```

Properties added in this fashion will not propagate to any other objects:

```
>>> do.granularity
Traceback (most recent call last):
    ...
AttributeError: 'Drug' object has no attribute 'granularity'
```

They will, however, be saved along with the object they are attached to.

2.5 Sharing Data with other users

Sharing is key to owmeta. This document covers the appropriate way to share changes with other owmeta users.

The shared owmeta database is stored in a Git repository distinct from the owmeta source code. Currently the database is stored in a Github repository here .

When you create a database normally, it will be stored in a format which is opaque to humans. In order to share your database you have two options: You can share the scripts which are used to create your database or you can share a human-readable serialization of the database. The second option is better since it doesn't require re-running your script to use the generated data, but it is best to share both.

For sharing the serialization, you should first clone the repository linked above, read the current serialization into your database (see *below* for an example of how you would do this), and then write out the serialization:

```
import owmeta as P
P.connect('path/to/your/config/file')
P.config()['rdf.graph'].serialize('out.n3', format='n3')
P.disconnect()
```

Commit, your changes to the git repository, push to a fork of the repository on Github and submit a pull request on the main repository. If for some reason your are unwilling or unable to create a Github account, post to the OpenWorm-discuss mailing list with a patch on the main repository with your changes and someone will have a look, possibly ask for adjustments or justification for your addition, and ultimately merge the changes for you.

To read the database back in you would do something like:

```
import owmeta as P
P.connect('path/to/your/config/file')
P.config()['rdf.graph'].parse('out.n3', format='n3')
P.disconnect()
```

Scripts are also added to the repository on Github to the `scripts` subdirectory.

2.6 Working with contexts

Contexts are introduced in *Adding Data to YOUR OpenWorm Database*. Here we provide a little more... context.

2.6.1 Background

Contexts were introduced to owmeta as a generic tool for grouping statements. We need to group statements to make statements about statements like "Who made these statements?" or "When were these statements made?". That's the main usage. Beyond that, we need a way to share statements. Contexts have identifiers by which we can naturally refer to contexts from other contexts.

owmeta needs a way to represent contexts with the existing statement form. Other alternatives were considered, such as using Python's context managers, but I (Mark) also wanted a way to put statements in a context that could also be carried with the subject of the statement. Using the wrapt package's proxies allows to achieve this while keeping the interface of the wrapped object the same, which is useful since it doesn't require a user of the object to know anything about contexts unless they need to change the context of a statement.

The remainder of this page will go into doing some useful things with contexts.

2.6.2 Classes and contexts

owmeta can load classes as well as instances from an RDF graph. The packages which define the classes must already be installed in the Python library path, and a few statements need to be in the graph you are loading from or in a graph imported (transitively) by that graph. The statements you need are these

```
:a_class_desc <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <http://openworm.org/
↪entities/PythonClassDescription> .
:a_class_desc <http://openworm.org/entities/ClassDescription/module> :a_module .
:a_class_desc <http://openworm.org/entities/PythonClassDescription/name> "AClassName"␣
↪.

:a_module <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <http://openworm.org/
↪entities/PythonModule> .
:a_module <http://openworm.org/entities/PythonModule/name> "APackage.and.module.name"␣
↪.
```

where `:a_class_desc` and `:a_module` are placeholders for objects which will typically be created by owmeta on the user's behalf, and `AClassName` is the name of the class available at the top-level of the module `APackage.and.module.name`. These statements will be created in memory by owmeta when a module defining a *`DataObject`*-derived class is first processed by a *`Mapper`* which will happen after the module is imported.

2.7 `owm` Command Line

The `owm` command line provides a high-level interface for working with owmeta-managed data. The central object which `owm` works on is the repository, which contains the triple store – a set of files in a binary format. The sub-commands act on important files inside the repository or with entities in the database.

To get usage information:

```
owm --help
```

To clone a repository:

```
owm clone $database_url
```

This will clone a repository into `.owm` in your current working directory. After a successful clone, a binary database usable as a owmeta store will have been created from the serialized graphs in the repository.

To save changes made to the database:

```
owm commit -m "I'm a commit message!"
```

To recreate the database from serialized graphs, *including uncommited changes*:

```
owm regendb
```

To make a new repository:

```
owm init
```

This will create a repository in `.owm` in your current working directory.

2.8 Software Versioning

The owmeta library follows the semanitc versioning scheme. For the sake of versioning, the software interface consists of:

1. the `owm` command line interface,
2. the underlying *owmeta.command.OWM* class underlying that CLI,
3. all "public" definitions in the *owmeta* package, sub-packages, and sub-modules,
4. the format of RDF data generated by *owmeta.dataObject.DataObject* and the subclasses thereof defined in the *owmeta* package, sub-packages, and sub-modules,
5. the API documentation for the *owmeta* package, sub-packages, and sub-modules,

In addition, any changes to the packages released on PyPI mandates at least a patch version increment.

For Git, our software version control system, software releases will be represented as tags in the form `v$semantic_version` with all components of the semantic version represented.

2.8.1 Documentation versioning

The documentation will have a distinct version number from the software. The version numbers for the documentation must change at least as often as the software versioning since the relationship of the documentation to the software necessarily changes. However, changes _only_ to the non-API documentation will not be a cause for a change to any of the components of the software version number. For documentation releases which coincide with software releases, the documentation version number will simply be the software version number. Any subsequent change to documentation between software releases will compel an increase in the documentation version number by one. The documentation version number for such documentation releases will be represented as `${software_version}+docs${documentation_increment}`.

2.9 Python Release Compatibility

All Python releases will be supported until they reach their official end-of-life, typically reported as "Release Schedule" PEPs (search "release schedule" on the PEP index) Thereafter, any regressions due to dependencies of owmeta dropping support for an EOL Python version, or due to a change in owmeta making use of a feature in a still-supported Python release will only be fixed for the sake of OpenWorm projects when requested by an issue on our tracker or for other projects when a compelling case can be made.

This policy is intended to provide support to most well-maintained projects which depend on owmeta while not overburdening developers.

2.10 BitTorrent client for P2P filesharing

1. **Download** desired contents:

- A *LocalFileDataSource* created and stored within the local graph store contains a torrent_file_name Informational. This refers to the torrent containing the location of the desired contents on the BitTorrent. A torrent is used to locate files on the File System [BEP 3]. A *DataSource* defines attributes about the contents that it represents.

Module t describes the DataSource attributes:

```
def owm_data(ns):
    ns.context.add_import(ConnectomeCSVDataSource.definition_context)
    ns.context(ConnectomeCSVDataSource)(
    key = '2000_connections',
    csv_file_name = 'connectome.csv',
    torrent_file_name = 'd9da5ce947c6f1c127dfcdc2ede63320.torrent'
 )
```

The DataSource can be created and stored on the local graph with:

```
$ owm save t
```

The DataSource identifier can be used to see contents stored in the local graph with:

```
$ owm source show ConnectomeCSVDataSource:2000_connections
```

ConnectomeCSVDataSource CSV file name: 'connectome.csv'

File name: 'connectome.csv'

Torrent file name: 'd9da5ce947c6f1c127dfcdc2ede63320.torrent'

- The BitTorrentDataSourceDirLoader class inherits from the *DataSourceDirLoader* and overrides its `load()` method. Google Drive stores the `torrents` uploaded by other researchers. `load()` fetches the `torrent` refered to in `torrent_file_name` of the `DataSource`, performs *DataTranslator* from one form to another and then adds the `torrent` to the BitTorrent Client for downloading its contents.

  This `BitTorrent Client` is available on PyPI and is included in the owmeta setup.

  To install separately:

```
$ pip install torrent-client
```

  For reference, use the torrent-client repository and its usage information with:

```
$ torrent_cli.py -h
```

  The `DataSourceDirLoader` attribute - `base_directory`, which is set in the `BitTorrentDataSourceDirLoader` constructor is where both the `torrent` and its contents are downloaded:

```
content = BitTorrentDataSourceDirLoader("./")
```

- Within the .owm directory we have the credentials.json and token.pickle these are for authentication of the Google Drive. For the purpose of access control the `client_secret` required by `credentials.json` will only be shared by owmeta maintainers.

- The `torrent` file name is the MD5 message digest of its contents. If the hash of the downloaded contents is the same as its `torrent` name the data is unaltered.

  Data-Integrity is to be checked after 100% download completion:

```
$ python3 integrity.py 'd9da5ce947c6f1c127dfcdc2ede63320.torrent' 'Merged_
↪Nuclei_Stained_Worm.zip'
```

2. **Upload** your contents:

- On an AWS EC2 instance is running a Nginx WSGI and a Flask Server to accept .zip content file uploads. Visit this Elastic IP address [13.235.204.78] to upload your files by browsing through your filesystem and then clicking the `Submit Query button`.

- This will create a `torrent` and `seed` your contents in parts, to other peers on the BitTorrent network. Content can then be downloaded as described above.

CHAPTER 3

For Developers

3.1 Testing in owmeta

3.1.1 Preparing for tests

owmeta should be installed like:

```
python setup.py develop
```

The default database should be populated like:

```
owm clone https://github.com/openworm/OpenWormData.git
```

3.1.2 Running tests

Tests should be run via setup.py like:

```
python setup.py test
```

you can pass options to `pytest` like so:

```
python setup.py test --addopts '-k DataIntegrityTest'
```

3.1.3 Writing tests

Tests are written using Python's unittest. In general, a collection of closely related tests should be in one file. For selecting different classes of tests, tests can also be tagged using pytest marks like:

```
@pytest.mark.tag
class TestClass(unittest.TestCase):
    ...
```

Currently, marks are used to distinguish between unit-level tests and others which have the `inttest` mark

3.2 Adding documentation

Documentation for owmeta is housed in two locations:

1. In the top-level project directory as `INSTALL.md` and `README.md`.
2. As a Sphinx project under the `docs` directory

By way of example, to add a page about useful facts concerning C. elegans to the documentation, include an entry in the list under `toctree` in `docs/index.rst` like:

```
worm-facts
```

and create the file `worm-facts.rst` under the `docs` directory and add a line:

```
.. _worm-facts:
```

to the top of your file, remembering to leave an empty line before adding all of your wonderful worm facts.

You can get a preview of what your documentation will look like when it is published by running `sphinx-build` on the docs directory:

```
sphinx-build -w sphinx-errors docs build_destination
```

The docs will be compiled to html which you can view by pointing your web browser at `build_destination/index.html`. If you want to view the documentation locally with the ReadTheDocs theme you'll need to download and install it.

3.2.1 API Documentation

API documentation is generated by the Sphinx autodoc extension. The format should be easy to pick up on, but a reference is available here. Just add a docstring to your function/class/method and add an `automodule` line to `owmeta/__init__.py` and your class should appear among the other documented classes.

3.2.2 Substitutions

Project-wide substitutions can be (conservatively!) added to allow for easily changing a value over all of the documentation. Currently defined substitutions can be found in `conf.py` in the `rst_epilog` setting. More about substitutions

3.2.3 Conventions

If you'd like to add a convention, list it here and start using it. It can be reviewed as part of a pull request.

1. Narrative text should be wrapped at 80 characters.
2. Long links should be extracted from narrative text. Use your judgement on what 'long' is, but if it causes the line width to stray beyond 80 characters that's a good indication.

3.3 RDF semantics for owmeta

In the context of owmeta, biological objects are classes of, for instance, anatomical features of a worm. That is to say, statements made about C. elegans are not about a specific worm, but are stated about the entire class of worms. The semantics of a `property SimpleProperty/value value` triple are that if any value is set, then without any additional statements being made, an instance of the object has been observed to have the value at some point in time, somewhere, under some set of conditions. In other words, the statement is an existential quantification over the associated object(class).

The purpose of the identifiers for Properties is to allow statements to be made about them directly. An example:

```
<http://openworm.org/entities/Entity/1> <http://openworm.org/entities/Entity/
↪interactsWith> <http://openworm.org/entities/Entity_interactsWith/2> .
<http://openworm.org/entities/Entity_interactsWith/2> <http://openworm.org/entities/
↪SimpleProperty/value> <http://openworm.org/entities/Entity/3> .

<http://openworm.org/entities/Entity/4> <http://openworm.org/entities/Entity/
↪modulates> <http://openworm.org/entities/Entity_modulates/5> .
<http://openworm.org/entities/Entity_modulates/5> <http://openworm.org/entities/
↪SimpleProperty/value> <http://openworm.org/entities/Entity_interactsWith/2>
```

3.4 RDF structure for owmeta

For most use cases, it is (hopefully) not necessary to write custom queries over the RDF graph in order to work with owmeta. However, if it does become necessary, it will be helpful to have an understanding of the structure of the RDF graph. Thus, a summary is given below.

For all *DataObjects* which are not *Properties*, there is an identifier of the form

```
<http://openworm.org/entities/Object_type/md5sum>
```

stored in the graph. This identifier will be associated with type data:

```
<http://openworm.org/entities/Object_type/md5sum> rdf:type <http://openworm.org/
↪entities/Object_type> .
<http://openworm.org/entities/Object_type/md5sum> rdf:type <http://openworm.org/
↪entities/parent_of_Object_type> .
<http://openworm.org/entities/Object_type/md5sum> rdf:type <http://openworm.org/
↪entities/parent_of_parent_of_Object_type> .
...
```

Properties have a slightly different form. They also have an identifier, which for `SimpleProperties` will look like this:

```
<http://openworm.org/entities/OwnerType_propertyName/md5sum>
```

`OwnerType` is the type of the Property's owner and `propertyName` is the name by which the property is accessed from an object of the owner's type. Other Properties will not necessarily have this form, but all of the standard Properties are implemented in terms of SimpleProperties and have no direct representation in the graph. For other Properties it is necessary to refer to their documentation or to examine the triples released by the Property of interest.

A DataObject's identifier is connected to a property in a triple like:

```
<http://openworm.org/entities/OwnerType/md5sum> <http://openworm.org/entities/
↪OwnerType/propertyName> <http://openworm.org/entities/OwnerType_propertyName/md5sum>
```

and the property is connected to its values like:

```
<http://openworm.org/entities/OwnerType_propertyName/md5sum> <http://openworm.org/
↪entities/SimpleProperty/value> "A literal value"
```

The following API calls do not yet exist, but would be excellent next functions to implement

3.5 Population()

A collection of cells. Constructor creates an empty population.

3.5.1 Population.filterCells(filters : ListOf(PairOf(unboundMethod, methodArgument))) : Population

Allows for groups of cells to be created based on shared properties including neurotransmitter, anatomical location or region, cell type.

Example:

```
p = Worm.cells()
p1 = p.filterCells([(Cell.lineageName, "AB")]) # A population of cells with AB as the␣
↪blast cell
```

3.6 NeuroML()

A utility for generating NeuroML files from other objects. The semantics described *above* do not apply here.

3.6.1 NeuroML.generate(object : {Network, Neuron, IonChannel}, type : {0,1,2}) : neuroml.NeuroMLDocument

Get a NeuroML object that represents the given object. The `type` determines what content is included in the NeuroML object:

- 0=full morphology+biophysics
- 1=cell body only+biophysics
- 2=full morphology only

3.6.2 NeuroML.write(document : neuroml.NeuroMLDocument, filename : String)

Write out a NeuroMLDocument

3.7 owmeta coding standards

Pull requests are *required* to follow the PEP-8 Guidelines for contributions of Python code to owmeta, with some exceptions noted below. Compliance can be checked with the `pep8` tool and these command line arguments:

```
--max-line-length=120 --ignore=E261,E266,E265,E402,E121,E123,E126,E226,E24,E704,E128
```

Refer to the pep8 documentation for the meanings of these error codes.

Lines of code should only be wrapped before 120 chars for readability. Comments and string literals, including docstrings, can be wrapped to a shorter length.

Some violations can be corrected with `autopep8`.

3.8 Design documents

These comprise the core design artifacts for owmeta.

3.8.1 Project Bundles

Project bundles are:

- collections of contexts,
- and a set of mappings between project-scoped human-friendly names and context identifiers.

They solve the problem of contexts containing different statements having the same identifier.

There are several ways we can get different contexts with the same identifier:

- through revisions of a context over time,
- by distinct groups using the same context identifier,
- or by contexts being distributed with different variants (e.g., a full and an abridged version).

In solving this problem of context ID aliasing, bundles also helps solve the problem of having contexts with inconsistent statements in the same project by providing a division within a project, between groups of contexts that aren't necessarily related.

Relationships

Where not specified, the subject of a relationship can participate in the relationship exactly once. For example, "A Dog has a Human", means "A Dog has one and only one Human"

- A Project can have zero or more Bundles
- A Bundle can belong to only one Project
- A Human-Friendly Name is associated with a Content-Based Identifier
- A Content-Based Identifier has one or more Hashes
- A Hash can appear in zero or more Content-Based Identifiers
- A Hash has an Algorithm ID and a Message Digest
- A Content-Based Identifier has an optional Tag
- There is at most one Content-Based Identifier for a given Tag

Types

Below is a description in terms of lower-level types of some higher-level types referenced above.

- A Tag is an arbitrary string
- A Message Digest is a Base-64 encoding of a string of bytes

3.8.2 Project Distribution

Projects are distributed as *bundle* archives, also referred to as `dists` (short for distributions) in the documentation and commands. The layout of files in a dist is essentially the same as the format of a `.owm` directory on initial clone. In other words the bundle contains a set of serialized graphs, an index of those graphs, an optional set of non-RDF data that accompanies data sources stored amongst the graphs, and a configuration file which serves as a working owmeta configuration file and a place for metadata about the bundle. The archive file format can be allowed to vary, between producers and consumers of dists, but at least the `tar.gz` format should be supported by general-purpose clients.

3.8.3 Data Packaging Lifecycle

The package lifecycle encompasses the creation of data, packaging of said data, and uploading to shared resources. The data packaging lifecycle draws from the Maven build lifecycle in the separation of local actions (e.g., `compile`, `stage`, `install` phases) from remote interactions (the `deploy` phase). To explain why we have these distinct phases, we should step back and look at what needs to happen when we share data.

In owmeta, we may be changing remote resources outside of the owmeta system. We also want to support local use and staging of data because it is expected that there is a lengthy period of data collection/generation, analysis, curation, and editing which precedes the publication of any data set. Having separate phases allows us to support a wider range of use-cases with owmeta in this local "staging" period.

To make the above more concrete, the prototypical example for us is around *`LocalFileDataSource`*, which wants to make the files described in the data source available for download. Typically, the local path to the file isn't useful outside of the machine. Also, except for files only a few tens of bytes in size, it isn't feasible to store the file contents in the same database as the metadata. We, still want to support metadata about these files and to avoid the necessity of making *n* different *`DataSource`* sub-classes for *n* different ways of getting a file. What we do is define a "deploy" phase that takes every *`LocalFileDataSource`* and "deploys" the files by uploading them to one or more remote stores or, in the case of a peer-to-peer solution, by publishing information about the file to a tracker or distributed hash table.

Packaging proceeds in phases to serve as an organizational structure for data producers, software developers, management, and information technology personnel. Compared with a more free-form build strategy like using an amalgam of shell scripts and disconnected commands, or even rule-based execution (e.g., GNU make), phases organize the otherwise implicit process by which the local database gets made available to other people. This explicitness is very useful since, when different people can take different roles in creating the configuration for each phase, having named phases where things happen aids in discussion, process development, and review. For instance, junior lab technicians may be responsible for creating or maintaining packaging with guidance from senior technicians or principal investigators. IT personnel may be interested in all phases since they all deal with the computing resources they manage, but they may focus on the phases that affect "remote" resources since those resources may, in fact, be managed within the same organization and require more effort in sharing URLs, generating access credentials, etc.

The remainder of this document will describe the default lifecycle and what takes place within each phase.

Default Lifecycle

The default lifecycle takes a *bundle*, including the contents of a owmeta triple store, creates one or more packages from that, stages the packages for ongoing development, and, finally, deploys packages to shared resources so that

colleagues and other interested parties can access them. Each phase is associated with a sub-command in *owm*.

Stage

Preparation for distribution.

When we're generating data, our workspace is not necessarily in the right state for distribution. We may have created temporary files and notes to ourselves, or we may have generated data in trial runs (or by mistake) which do not reflect our formal experimental conditions. In the staging phase, we bring together just the data which we wish to distribute for a given bundle. During the staging phase we also serialize

Once these data are brought together in the staging area, they should be immutable – in other words, they should not change any more. Consequently, the staging phase is the appropriate time for creating summary statistics, signatures, and content-based identifiers.

For files associated with staged RDF (Resource Description Framework) data, Much of the data which is created in a research lab is append-only: observations are logged and timestamped either by a human or by a machine in the moment they happen, but, if done properly, such logs are rarely edited, or, if there is an amendment, it also is logged as such, with the original record preserved. As long as this append-only property is preserved, we only need to designate the range of such time-stamped records which belong in a package to have the desired immutability. Of course, if the source data is expected to be changed, then we would want either a copy-on-write mechanism (at the file system level) or to copy the files. Regardless, file hashes and/or signatures created during the staging phase would be available for guarding against accidental changes.

Install

Local installation. Preparation for deployment.

The "install" phase takes the staged data, and adds additional glue to make it available on the local machine as it would be for a remote machine after deployment. owmeta will create a local repository to house installed packages. The repository stores the relationship between the human-friendly name for the package (serving a purpose similar to Maven's group-artifact-version coordinates) and the set of serialized RDF graphs in the package. Given that the repository is meant to serve a user across projects, the repository will be stored in the "user directory", if one can be found on the system.[1]

Continuing the pattern of putting configuration in RDF form, the repository is also described as in RDF and shall use the same form as remote repositories, up to a substitution of access protocols (e.g., file system access in place of HTTP access). The value here is in interoperability and ease of implementation. For the first point, we have fairly broad support for RDF query and manipulation across programming languages. The second point is supported by the first and by the fact that, once we've got implementations for the necessary access methods, no additional code should need to be written for access to remote repositories beyond what's done for local.

The same argument about immutability of data files applies to the install phase as well. Installed packages may still have references to paths on the local file system. It is not until the deploy phase that all local paths must be expunged.

Deploy

Creation of configuration for upload/download. Sharing packages.

In the "deploy" phase, we publish our data to "remotes". A "remote" may be a repository or, in the case of a peer-to-peer file sharing system, a file index or DHT. Above, we referred to non-RDF data files on the local file system – during the deploy phase, these files are actually published and accession information (e.g., a database record identifier) for those files is generated and returned to the system where the deployment was initiated. This assumes a fully automated

[1] This will be the user directory as determined by `os.path.expanduser()`

process for publication of files: If, instead, the publication platform requires some manual interaction, that must be done outside of owmeta and then the accession information would be provided with the deploy command.

3.8.4 Publishing DataSources

DataSource is a subclass of *DataObject* with a few features to make describing data files (CSV, HDF5, Excel) a bit more consistent and to make recovering those files, and information about them, more reliable. In order to have that reliability we have to take some extra measures when publishing a `DataSource`. In particular, we must publish local files referred to by the `DataSource` and relativize those references. This file publication happens in the *"deploy" phase* of the data packaging lifecycle. Before that, however, a description of what files need to be published is generated in the *"stage" phase*. In the "stage" phase, the `DataSources` with files needing publication are queried for in the configured triple store, and the "staging manager", the component responsible for coordinating the "stage" phase identifies file references that refer to the same files and directories.

3.9 Querying for data objects

3.9.1 X.query form

Creates modified version of the DataObject subclass which is fit for using in queries. May do other additional things latter, but, principally, it overrides the identifier generation based on attributes.

Examples for querying for a *Neuron* object:

```
Neuron.query(name='AVAL')
ctx(Neuron).query(name='AVAL')
ctx.stored(Neuron).query(name='AVAL')
ctx.mixed(Neuron).query(name='AVAL')
```

this can be important for when a class generates identifiers based on some number of properties, but a subclass doesn't use the superclass identifier scheme (*Cell* and Neuron are an example). The query form allows to query from the superclass as you normally would to get subclass instances.

CHAPTER 4

Issues

CHAPTER 5

Indices and tables

- genindex
- modindex
- search

Python Module Index

o

Index

D

E

F

G

H

I

L

M

N

O

P

Q

R

S

T

U

V

W

X

Y

Z