
TranSMART loader Documentation

Release 0.1.3

Gijs Kant

May 15, 2019

Contents:

1	Installation and usage	3
1.1	Usage	3
2	Documentation	5
3	Known issues	7
4	Development	9
4.1	Python versions	9
4.2	Package management and dependencies	9
4.3	Testing and code coverage	9
4.4	Documentation	10
4.5	Coding style conventions and code quality	10
5	License	11
6	Credits	13
7	Data model	15
7.1	Data model	15
8	API Reference	19
8.1	transmart_loader package	19
9	Index	27
	Python Module Index	29

This package contains classes that represent the core domain objects stored in the [TranSMART](#) platform, an open source data sharing and analytics platform for translational biomedical research.

It also provides a utility that writes such objects to tab-separated files that can be loaded into a TranSMART database using the [transmart-copy](#) tool.

Note: this is a very preliminary version, still under development. Issues can be reported at https://github.com/thehyve/python_transmart_loader/issues.

CHAPTER 1

Installation and usage

To install transmart_loader, do:

```
pip install transmart-loader
```

or from sources:

```
git clone https://github.com/thehyve/python_transmart_loader.git
cd python_transmart_loader
pip install .
```

1.1 Usage

Usage examples can be found in these projects:

- [fhir2transmart](#): a tool that translates core [HL7 FHIR](#) resources to the TranSMART data model.
- [ontology2transmart](#): a tool that translates ontologies available from [DIMDI](#) to TranSMART ontologies.

CHAPTER 2

Documentation

Full documentation of the package is available at [Read the Docs](#).

CHAPTER 3

Known issues

- Date values are not correctly translated

CHAPTER 4

Development

For a quick reference on software development, we refer to [the software guide checklist](#).

4.1 Python versions

This repository is set up with Python version 3.6

Add or remove Python versions based on project requirements. [The guide](#) contains more information about Python versions and writing Python 2 and 3 compatible code.

4.2 Package management and dependencies

This project uses *pip* for installing dependencies and package management.

- Dependencies should be added to *setup.py* in the *install_requires* list.

4.3 Testing and code coverage

- Tests are in the `tests` folder.
- The `tests` folder contains:
 - A test if files for *transmart-copy* are generated for fake data (file: `test_transmart_loader`)
 - A test that checks whether your code conforms to the Python style guide (PEP 8) (file: `test_lint.py`)
- The testing framework used is [PyTest](#)
 - [PyTest introduction](#)
- Tests can be run with `python setup.py test`

4.4 Documentation

- Documentation should be put in the `docs` folder.
- To generate html documentation run `python setup.py build_sphinx`

4.5 Coding style conventions and code quality

- Check your code style with `prospector`
- You may need run `pip install .[dev]` first, to install the required dependencies

CHAPTER 5

License

Copyright (c) 2019 The Hyve B.V.

The TranSMART loader is licensed under the MIT License. See the file [LICENSE](#).

CHAPTER 6

Credits

This package was created with [Cookiecutter](#) and the [NLeSC/python-template](#).

CHAPTER 7

Data model

7.1 Data model

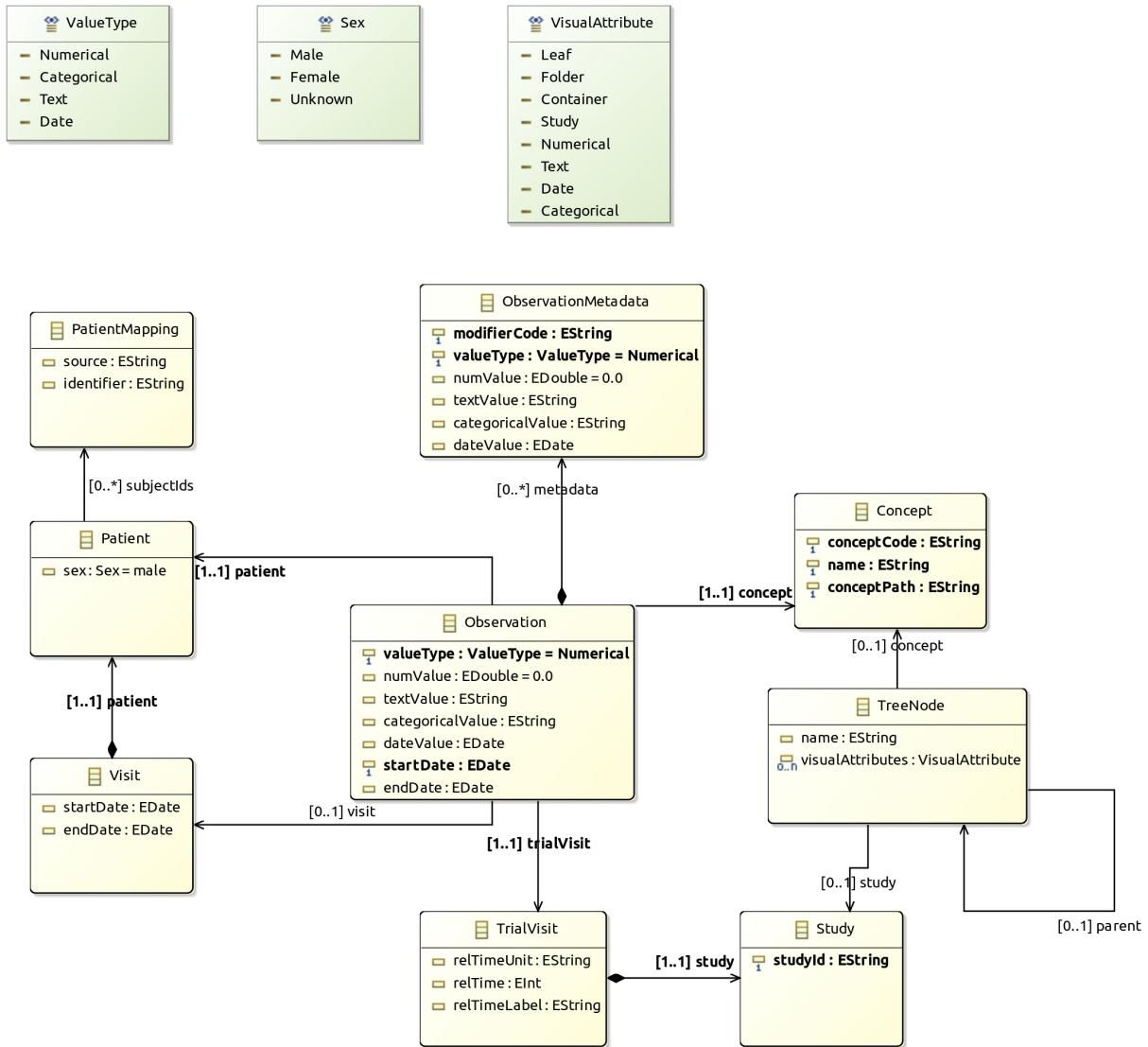


Fig. 1: TranSMART data model

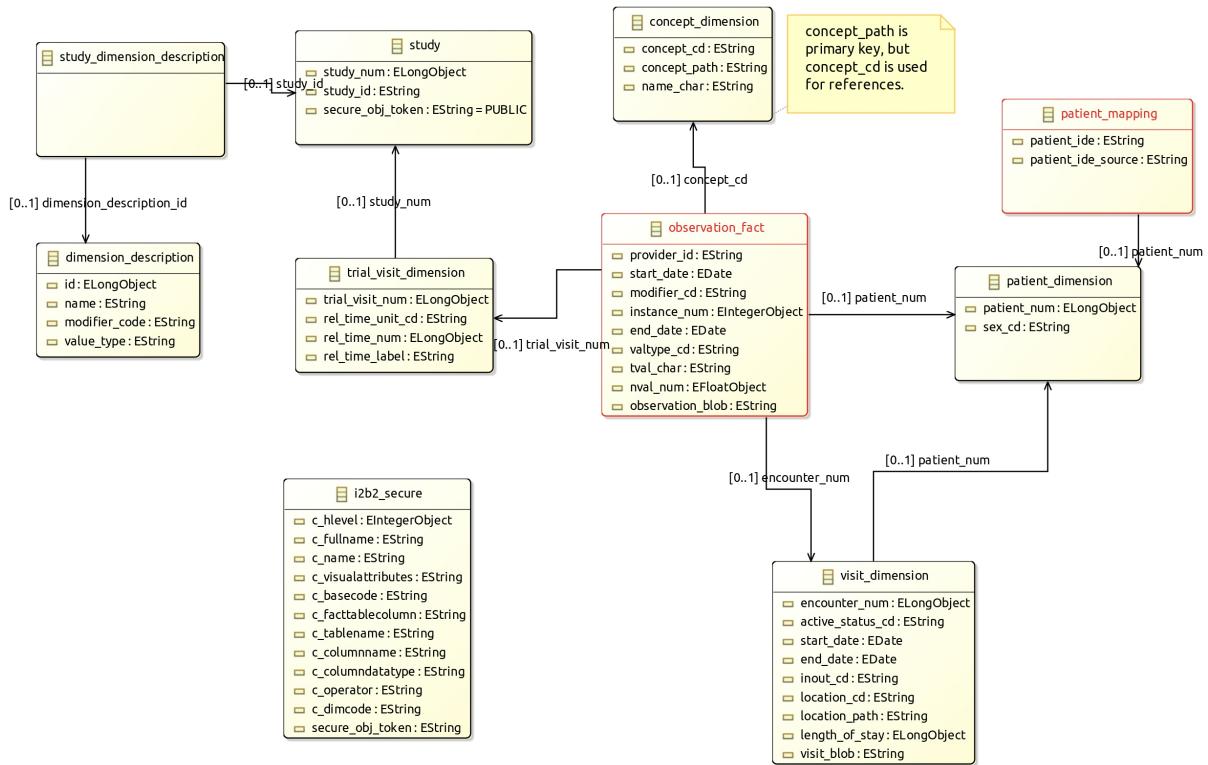


Fig. 2: TranSMART database tables

CHAPTER 8

API Reference

8.1 transmart_loader package

Documentation about TranSMART loader

8.1.1 Submodules

[transmart_loader.collection_validator module](#)

```
class transmart_loader.collection_validator.CollectionValidator
Bases: transmart_loader.collection_visitor.CollectionVisitor

Validation class for TranSMART data collections.

static validate(collection: transmart_loader.transmart.DataCollection)
visit_concept(concept: transmart_loader.transmart.Concept) → None
visit_node(node: transmart_loader.transmart.TreeNode) → None
visit_observation(observation: transmart_loader.transmart.Observation) → None
visit_patient(patient: transmart_loader.transmart.Patient) → None
visit_study(study: transmart_loader.transmart.Study) → None
visit_trial_visit(trial_visit: transmart_loader.transmart.TrialVisit) → None
visit_visit(visit: transmart_loader.transmart.Visit) → None
```

[transmart_loader.collection_visitor module](#)

```
class transmart_loader.collection_visitor.CollectionVisitor
Bases: object

Visitor class for TranSMART data collections
```

```
visit (collection: Optional[transmart_loader.transmart.DataCollection]) → None
visit_concept (concept: transmart_loader.transmart.Concept) → None
visit_node (node: transmart_loader.transmart.TreeNode) → None
visit_observation (observation: transmart_loader.transmart.Observation) → None
visit_patient (patient: transmart_loader.transmart.Patient) → None
visit_study (study: transmart_loader.transmart.Study) → None
visit_trial_visit (trial_visit: transmart_loader.transmart.TrialVisit) → None
visit_visit (visit: transmart_loader.transmart.Visit) → None
```

transmart_loader.console module

```
class transmart_loader.console.Console
Bases: object
```

A helper class for displaying messages on the console (stderr).

```
Black = '\x1b[30m'
BlackBackground = '\x1b[40m'
Blue = '\x1b[94m'
Green = '\x1b[92m'
GreenBackground = '\x1b[42m'
Grey = '\x1b[37m'
Red = '\x1b[91m'
RedBackground = '\x1b[41m'
Reset = '\x1b[0m'
Yellow = '\x1b[93m'
YellowBackground = '\x1b[103m'

static error(message)
static info(message)
static success(message)
static title(title)
static warning(message)
```

transmart_loader.copy_writer module

```
class transmart_loader.copy_writer.TransmartCopyWriter(output_dir: str)
Bases: transmart_loader.collection_visitor.CollectionVisitor
```

Writes TranSMART data collections to a folder with files that can be loaded into a TranSMART database using transmart-copy.

```
concepts_header = ['concept_cd', 'concept_path', 'name_char']
dimensions_header = ['id', 'name', 'modifier_code', 'value_type']
```

```
init_writers() → None
    Creates files and initialises writers for the output files in transmart-copy format.

observations_header = ['encounter_num', 'patient_num', 'concept_cd', 'provider_id', 's']
patient_mappings_header = ['patient_ide', 'patient_ide_source', 'patient_num']
patients_header = ['patient_num', 'sex_cd']

prepare_output_dir() → None
    Creates an output directory if it does not exist. Fails if the output directory exists and is not empty.

studies_header = ['study_num', 'study_id', 'secure_obj_token']
study_dimensions_header = ['study_id', 'dimension_description_id']
tree_nodes_header = ['c_hlevel', 'c_fullname', 'c_name', 'c_visualattributes', 'c_base'
trial_visits_header = ['trial_visit_num', 'study_num', 'rel_time_unit_cd', 'rel_time_n
value_type_codes = {<ValueType.Numeric: 1>: 'N', <ValueType.Categorical: 2>: 'T',
visit_concept (concept: transmart_loader.transmart.Concept) → None
    Serialises a Concept entity to a TSV file.

    Parameters concept – the Concept entity

visit_node (node: transmart_loader.transmart.TreeNode) → None

visit_observation (observation: transmart_loader.transmart.Observation) → None
    Serialises an Observation entity to a TSV file.

    FIXME: fix date value serialisation

    Parameters observation – the Observation entity

visit_patient (patient: transmart_loader.transmart.Patient) → None
    Serialises an Patient entity and related PatientMapping entities to TSV files.

    Parameters patient – the Patient entity

visit_study (study: transmart_loader.transmart.Study) → None
    Serialises a Study entity to a TSV file.

    Parameters study – the Study entity

visit_tree_node (node: transmart_loader.transmart.TreeNode, level=0, parent_path='\\')
    Serialises a TreeNode entity and its children to a TSV file.

    Parameters
        • node – the TreeNode entity
        • level – the hierarchy level of the node
        • parent_path – the path of the parent node.

visit_trial_visit (trial_visit: transmart_loader.transmart.TrialVisit) → None
    Serialises a TrialVisit entity to a TSV file.

    Parameters trial_visit – the TrialVisit entity

visit_visit (visit: transmart_loader.transmart.Visit) → None
    Serialises a Visit entity to a TSV file. NB: this requires all patient visits to be cleared before loading new
    visits for the patient.

    Parameters visit – the Visit entity
```

```
visits_header = ['encounter_num', 'patient_num', 'active_status_cd', 'start_date', 'end_date']
write_collection(collection: transmart_loader.transmart.DataCollection) → None
write_dimension(dimension: transmart_loader.transmart.Dimension) → None
    Serialises a Dimension entity to a TSV file.

    Parameters dimension – the Dimension entity

write_dimensions() → None
    Write dimensions metadata and link all studies to the dimensions

write_study_dimensions(study_index)

class transmart_loader.copy_writer.VisualAttribute
    Bases: enum.Enum

    Visual attribute of an ontology node

    Categorical = 8
    Container = 3
    Date = 7
    Folder = 2
    Leaf = 1
    Numerical = 5
    Study = 4
    Text = 6

transmart_loader.copy_writer.format_date(value: Optional[datetime.date]) → Optional[str]
transmart_loader.copy_writer.get_concept_node_row(node: transmart_loader.transmart.ConceptNode, level, node_path)
transmart_loader.copy_writer.get_folder_node_row(node: transmart_loader.transmart.TreeNode, level, node_path)
transmart_loader.copy_writer.get_study_node_row(node: transmart_loader.transmart.StudyNode, level, node_path)
```

transmart_loader.csv_types module

```
class transmart_loader.csv_types.CsvWriter
    Bases: object

    This is an abstract version of the return of csv.writer().

    writerow(row: Sequence[Any]) → None
    writerows(rows: Sequence[Sequence[Any]]) → None
```

transmart_loader.loader_exception module

```
exception transmart_loader.loader_exception.LoaderException
    Bases: Exception
```

transmart_loader.transmart module

```
class transmart_loader.transmart.CategoricalValue (value: str)
    Bases: transmart_loader.transmart.Value

    A categorical value

    value()

    value_type()

class transmart_loader.transmart.Concept (concept_code: str, name: str, concept_path: str, value_type: transmart_loader.transmart.ValueType)
    Bases: object

    Concepts to classify observations

class transmart_loader.transmart.ConceptNode (concept: transmart_loader.transmart.Concept)
    Bases: transmart_loader.transmart.TreeNode

    Concept node

class transmart_loader.transmart.DataCollection (concepts: Iterable[transmart_loader.transmart.Concept], studies: Iterable[transmart_loader.transmart.Study], trial_visits: Iterable[transmart_loader.transmart.TrialVisit], visits: Iterable[transmart_loader.transmart.Visit], ontology: Iterable[transmart_loader.transmart.TreeNode], patients: Iterable[transmart_loader.transmart.Patient], observations: Iterable[transmart_loader.transmart.Observation])
    Bases: object

    A data collection that can be loaded into TranSMART

class transmart_loader.transmart.DateValue (value: datetime.date)
    Bases: transmart_loader.transmart.Value

    A date value

    value()

    value_type()

class transmart_loader.transmart.Dimension (name: str, modifier_code: Optional[str] = None, value_type: Optional[transmart_loader.transmart.ValueType] = None)
    Bases: object

    Dimension metadata

class transmart_loader.transmart.NumericalValue (value: float)
    Bases: transmart_loader.transmart.Value

    A numerical value
```

```
    value()
    value_type()

class transmart_loader.transmart.Observation(patient: trans-
    mart_loader.transmart.Patient,
    concept: trans-
    mart_loader.transmart.Concept, visit: Op-
    tional[transmart_loader.transmart.Visit],
    trial_visit: trans-
    mart_loader.transmart.TrialVisit,
    start_date: Optional[datetime.date],
    end_date: Optional[datetime.date], value:
    transmart_loader.transmart.Value)

Bases: object
Data about an observed event or an attribute of a patient

class transmart_loader.transmart.ObservationMetadata
Bases: object
Metadata about an observation

class transmart_loader.transmart.Patient(identifier: str, sex: str, mappings: Se-
    quence[transmart_loader.transmart.PatientMapping])
Bases: object
Patient properties

class transmart_loader.transmart.PatientMapping(source: str, identifier: str)
Bases: object
Patient identifiers

class transmart_loader.transmart.Study(study_id: str, name: str)
Bases: object
Study node

class transmart_loader.transmart.TextValue(value: str)
Bases: transmart_loader.transmart.Value
A text value

    value()
    value_type()

class transmart_loader.transmartTreeNode(name: str)
Bases: object
Ontology node
    add_child(child: transmart_loader.transmartTreeNode)

class transmart_loader.transmart.TrialVisit(study: transmart_loader.transmart.Study,
    rel_time_label: str, rel_time_unit: Op-
    tional[str] = None, rel_time: Optional[int]
    = None)

Bases: object
Trial visit
```

```

class transmart_loader.transmart.Value
Bases: object

An observed value

value

value_type

class transmart_loader.transmart.ValueType
Bases: enum.Enum

Type of an observed value

Categorical = 2
Date = 4
Numeric = 1
Text = 3

class transmart_loader.transmart.Visit (patient: transmart_loader.transmart.Patient,
                                         identifier: str, active_status: Optional[str],
                                         start_date: Optional[datetime.date], end_date:
                                         Optional[datetime.date], inout: Optional[str], location:
                                         Optional[str], length_of_stay: Optional[int])
Bases: object

Patient visit

```

[transmart_loader.tsv_writer module](#)

```

class transmart_loader.tsv_writer.TsvWriter (path: str)
Bases: transmart_loader.csv_types.CsvWriter

Tab-separated values writer. Creates a new file when initialised and fails when the file already exists.

close () → None
writerow (row: Sequence[Any]) → None
writerows (rows: Sequence[Sequence[Any]]) → None

```


CHAPTER 9

Index

- genindex

Python Module Index

t

transmart_loader, 19
transmart_loader.collection_validator,
 19
transmart_loader.collection_visitor, 19
transmart_loader.console, 20
transmart_loader.copy_writer, 20
transmart_loader.csv_types, 22
transmart_loader.loader_exception, 22
transmart_loader.transmart, 23
transmart_loader.tsv_writer, 25

Index

A

add_child() (*transmart_loader.transmart.TreeNode* method), 24

B

Black (*transmart_loader.console.Console* attribute), 20
BlackBackground (*transmart_loader.console.Console* attribute), 20

Blue (*transmart_loader.console.Console* attribute), 20

C

Categorical (*transmart_loader.copy_writer.VisualAttribute* attribute), 22

Categorical (*transmart_loader.transmart.ValueType* attribute), 25

CategoricalValue (class in *transmart_loader.transmart*), 23

close() (*transmart_loader.tsv_writer.TsvWriter* method), 25

CollectionValidator (class in *transmart_loader.collection_validator*), 19

CollectionVisitor (class in *transmart_loader.collection_visitor*), 19

Concept (class in *transmart_loader.transmart*), 23

ConceptNode (class in *transmart_loader.transmart*), 23

concepts_header (*transmart_loader.copy_writer.TransmartCopyWriter* attribute), 20

Console (class in *transmart_loader.console*), 20

Container (*transmart_loader.copy_writer.VisualAttribute* attribute), 22

CsvWriter (class in *transmart_loader.csv_types*), 22

D

DataCollection (class in *transmart_loader.transmart*), 23

Date (*transmart_loader.copy_writer.VisualAttribute* attribute), 22

Date (*transmart_loader.transmart.ValueType* attribute), 25

DateValue (class in *transmart_loader.transmart*), 23

Dimension (class in *transmart_loader.transmart*), 23
dimensions_header (*transmart_loader.copy_writer.TransmartCopyWriter* attribute), 20

E

error() (*transmart_loader.console.Console* static method), 20

F

Folder (*transmart_loader.copy_writer.VisualAttribute* attribute), 22

format_date() (in module *transmart_loader.copy_writer*), 22

G

get_concept_node_row() (in module *transmart_loader.copy_writer*), 22

get_folder_node_row() (in module *transmart_loader.copy_writer*), 22

get_study_node_row() (in module *transmart_loader.copy_writer*), 22

Green (*transmart_loader.console.Console* attribute), 20

GreenBackground (*transmart_loader.console.Console* attribute), 20

Grey (*transmart_loader.console.Console* attribute), 20

info() (*transmart_loader.console.Console* static method), 20

init_writers() (*transmart_loader.copy_writer.TransmartCopyWriter* method), 20

I

L

Leaf (*transmart_loader.copy_writer.VisualAttribute attribute*), 22
LoaderException, 22

N

Numeric (*transmart_loader.transmart.ValueType attribute*), 25
Numerical (*transmart_loader.copy_writer.VisualAttribute attribute*), 22
NumericalValue (class in *transmart_loader.transmart*), 23

O

Observation (class in *transmart_loader.transmart*), 24
ObservationMetadata (class in *transmart_loader.transmart*), 24
observations_header (*transmart_loader.copy_writer.TransmartCopyWriter attribute*), 21

P

Patient (class in *transmart_loader.transmart*), 24
patient_mappings_header (*transmart_loader.copy_writer.TransmartCopyWriter attribute*), 21
PatientMapping (class in *transmart_loader.transmart*), 24
patients_header (*transmart_loader.copy_writer.TransmartCopyWriter attribute*), 21
prepare_output_dir () (*transmart_loader.copy_writer.TransmartCopyWriter method*), 21

R

Red (*transmart_loader.console.Console attribute*), 20
RedBackground (*transmart_loader.console.Console attribute*), 20
Reset (*transmart_loader.console.Console attribute*), 20

S

studies_header (*transmart_loader.copy_writer.TransmartCopyWriter attribute*), 21
Study (class in *transmart_loader.transmart*), 24
Study (*transmart_loader.copy_writer.VisualAttribute attribute*), 22
study_dimensions_header (*transmart_loader.copy_writer.TransmartCopyWriter attribute*), 21
StudyNode (class in *transmart_loader.transmart*), 24

success () (*transmart_loader.console.Console static method*), 20

T

Text (*transmart_loader.copy_writer.VisualAttribute attribute*), 22
Text (*transmart_loader.transmart.ValueType attribute*), 25

TextValue (class in *transmart_loader.transmart*), 24
title () (*transmart_loader.console.Console static method*), 20
transmart_loader (module), 19
transmart_loader.collection_validator (module), 19

transmart_loader.collection_visitor (module), 19

transmart_loader.console (module), 20
transmart_loader.copy_writer (module), 20
transmart_loader.csv_types (module), 22
transmart_loader.loader_exception (module), 22

transmart_loader.transmart (module), 23

transmart_loader.tsv_writer (module), 25

TransmartCopyWriter (class in *transmart_loader.copy_writer*), 20
tree_nodes_header (*transmart_loader.copy_writer.TransmartCopyWriter attribute*), 21

TreeNode (class in *transmart_loader.transmart*), 24

trial_visits_header (*transmart_loader.copy_writer.TransmartCopyWriter attribute*), 21

TrialVisit (class in *transmart_loader.transmart*), 24

TsvWriter (class in *transmart_loader.tsv_writer*), 25

V

validate () (*transmart_loader.collection_validator.CollectionValidator static method*), 19

Value (class in *transmart_loader.transmart*), 24

value (*transmart_loader.transmart.Value attribute*), 25

value () (*transmart_loader.transmart.CategoricalValue method*), 23

value () (*transmart_loader.transmart.DateValue method*), 23

value () (*transmart_loader.transmart.NumericalValue method*), 23

value () (*transmart_loader.transmart.TextValue method*), 24

value_type (*transmart_loader.transmart.Value attribute*), 25

value_type () (*transmart_loader.transmart.CategoricalValue method*), 23

value_type()	(trans-	visit_study()	(trans-
mart_loader.transmart.DateValue	method),	mart_loader.collection_visitor.CollectionVisitor	method), 20
23			
value_type()	(trans-	visit_study()	(trans-
mart_loader.transmart.NumericalValue	method), 24	mart_loader.copy_writer.TransmartCopyWriter	method), 21
value_type() (transmart_loader.transmart.TextValue	(trans-	visit_tree_node()	(trans-
method), 24		mart_loader.copy_writer.TransmartCopyWriter	method), 21
value_type_codes	(trans-	visit_trial_visit()	(trans-
mart_loader.copy_writer.TransmartCopyWriter	attribute), 21	mart_loader.collection_validator.CollectionValidator	method), 19
ValueType (class in transmart_loader.transmart), 25			
Visit (class in transmart_loader.transmart), 25		visit_trial_visit()	(trans-
visit() (transmart_loader.collection_visitor.CollectionVisitor		mart_loader.collection_visitor.CollectionVisitor	method), 20
method), 19			
visit_concept()	(trans-	visit_trial_visit()	(trans-
mart_loader.collection_validator.CollectionValidator	method), 19	mart_loader.copy_writer.TransmartCopyWriter	method), 21
visit_concept()	(trans-	visit_visit()	(trans-
mart_loader.collection_visitor.CollectionVisitor	method), 20	mart_loader.collection_validator.CollectionValidator	method), 19
visit_concept()	(trans-	visit_visit()	(trans-
mart_loader.copy_writer.TransmartCopyWriter	method), 21	mart_loader.collection_visitor.CollectionVisitor	method), 20
visit_node()	(trans-	visit_visit()	(trans-
mart_loader.collection_validator.CollectionValidator	method), 19	mart_loader.copy_writer.TransmartCopyWriter	method), 21
visit_node()	(trans-	visits_header	(trans-
mart_loader.collection_visitor.CollectionVisitor	method), 20	mart_loader.copy_writer.TransmartCopyWriter	attribute), 21
visit_node()	(trans-	VisualAttribute (class in	trans-
mart_loader.copy_writer.TransmartCopyWriter	method), 21	mart_loader.copy_writer), 22	
visit_observation()	(trans-	W	
mart_loader.collection_validator.CollectionValidator	method), 19	warning() (transmart_loader.console.Console static	
visit_observation()	(trans-	method), 20	
mart_loader.collection_visitor.CollectionVisitor	method), 20	write_collection()	(trans-
visit_observation()	(trans-	mart_loader.copy_writer.TransmartCopyWriter	method), 22
mart_loader.copy_writer.TransmartCopyWriter	method), 21	method), 22	
visit_patient()	(trans-	write_dimension()	(trans-
mart_loader.collection_validator.CollectionValidator	method), 19	mart_loader.copy_writer.TransmartCopyWriter	method), 22
visit_patient()	(trans-	method), 22	
mart_loader.collection_visitor.CollectionVisitor	method), 20	write_study_dimensions()	(trans-
visit_patient()	(trans-	mart_loader.copy_writer.TransmartCopyWriter	method), 22
mart_loader.copy_writer.TransmartCopyWriter	method), 21	method), 22	
visit_study()	(trans-	writerow() (transmart_loader.csv_types.CsvWriter	
mart_loader.collection_validator.CollectionValidator	method), 19	method), 22	
		writerow() (transmart_loader.tsv_writer.TsvWriter	
		method), 25	
		writerows() (transmart_loader.csv_types.CsvWriter	
		method), 22	

`writerows()` (*transmart_loader.tsv_writer.TsvWriter*
method), 25

Y

`Yellow` (*transmart_loader.console.Console* attribute),
20

`YellowBackground` (*trans-
mart_loader.console.Console* attribute),
20