OpenAlex Analysis

Generated by Doxygen 1.9.1

Chapter 1

Automatic literature analysis

Sustainability research references, network, methods and theory landscape analysis within the Stockholm Resilience Centre.

This repo provides classes and methods to extract statistics, plots and graphs, as well as examples in Jupyter Notebooks from the OpenAlex API. There is also a web app built with Dash and some docker config files to deploy it.

For the documentation, open `doc/html/index.html` (you need to download the repo to open it in the browser for now).

Romain Thomas 2023

Chapter 2

Hierarchical Index

2.1 Class Hierarchy

This inheritance list is sorted roughly, but not completely, alphabetically:

Chapter 3

Class Index

3.1 Class List

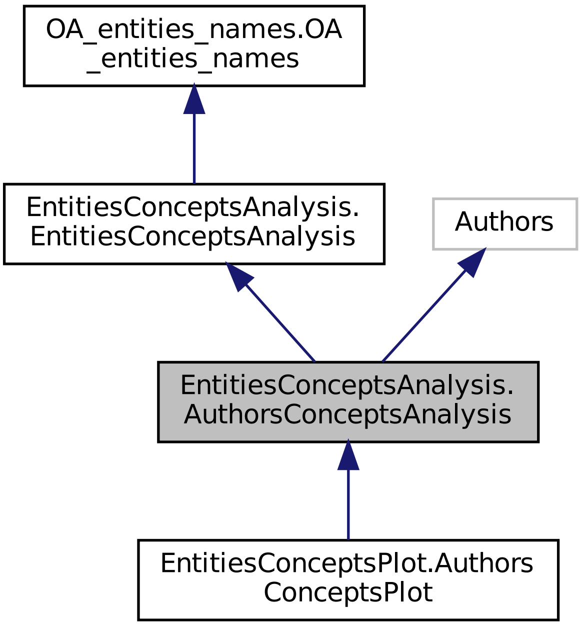
Here are the classes, structs, unions and interfaces with brief descriptions:

Chapter 4

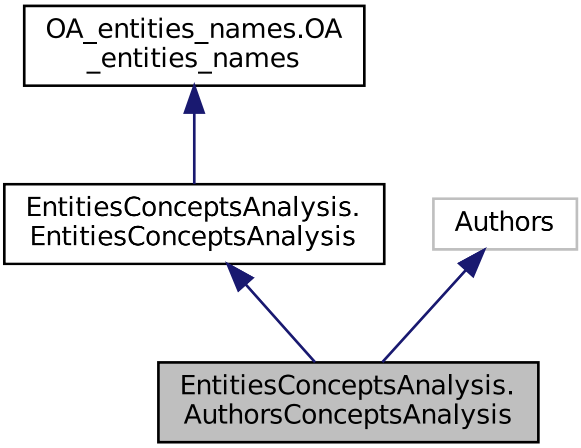
Class Documentation

4.1 EntitiesConceptsAnalysis.AuthorsConceptsAnalysis Class Reference

Inheritance diagram for EntitiesConceptsAnalysis.AuthorsConceptsAnalysis:



Collaboration diagram for EntitiesConceptsAnalysis.AuthorsConceptsAnalysis:



Static Public Attributes

- **EntitieOpenAlex** = Authors

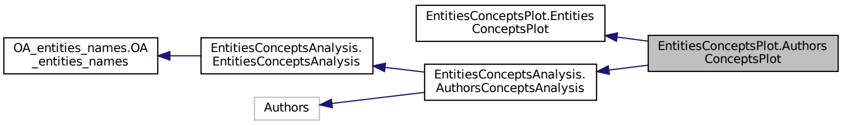
Additional Inherited Members

The documentation for this class was generated from the following file:

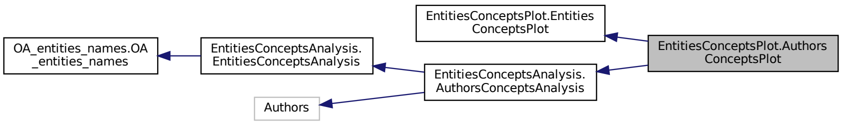
- EntitiesConceptsAnalysis.py

4.2 EntitiesConceptsPlot.AuthorsConceptsPlot Class Reference

Inheritance diagram for EntitiesConceptsPlot.AuthorsConceptsPlot:



Collaboration diagram for EntitiesConceptsPlot.AuthorsConceptsPlot:



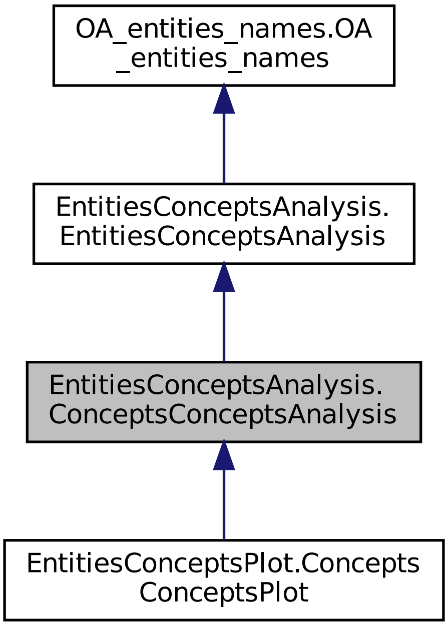
Additional Inherited Members

The documentation for this class was generated from the following file:

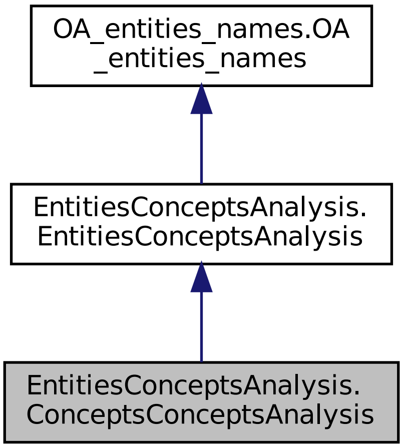
- EntitiesConceptsPlot.py

4.3 EntitiesConceptsAnalysis.ConceptsConceptsAnalysis Class Reference

Inheritance diagram for EntitiesConceptsAnalysis.ConceptsConceptsAnalysis:



Collaboration diagram for EntitiesConceptsAnalysis.ConceptsConceptsAnalysis:



Static Public Attributes

- **EntitieOpenAlex** = Concepts

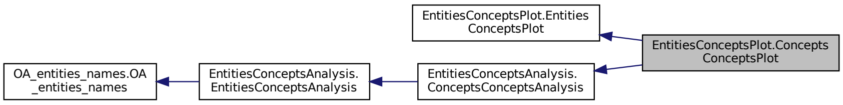
Additional Inherited Members

The documentation for this class was generated from the following file:

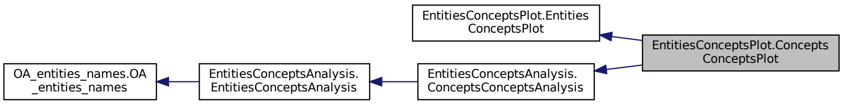
- EntitiesConceptsAnalysis.py

4.4 EntitiesConceptsPlot.ConceptsConceptsPlot Class Reference

Inheritance diagram for EntitiesConceptsPlot.ConceptsConceptsPlot:


EntitiesConceptsPlot.Entities
ConceptsPlot
EntitiesConceptsPlot.Concepts
ConceptsPlot
OA_entities_names.OA
_entities_names
EntitiesConceptsAnalysis.
EntitiesConceptsAnalysis
EntitiesConceptsAnalysis.
ConceptsConceptsAnalysis

Collaboration diagram for EntitiesConceptsPlot.ConceptsConceptsPlot:


EntitiesConceptsPlot.Entities
ConceptsPlot
EntitiesConceptsPlot.Concepts
ConceptsPlot
OA_entities_names.OA
_entities_names
EntitiesConceptsAnalysis.
EntitiesConceptsAnalysis
EntitiesConceptsAnalysis.
ConceptsConceptsAnalysis

Additional Inherited Members

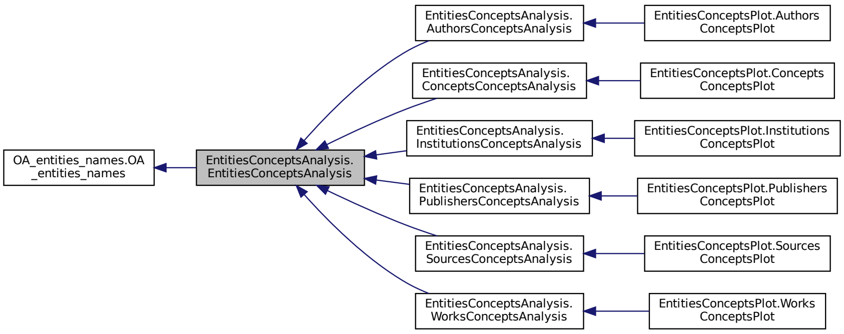
The documentation for this class was generated from the following file:

- EntitiesConceptsPlot.py

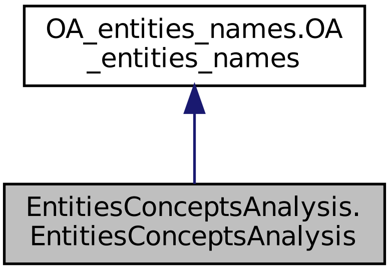
4.5 EntitiesConceptsAnalysis.EntitiesConceptsAnalysis Class Reference

OpenAlexAnalysis class which contains generic methods to do analysis over OpenAlex entities.

Inheritance diagram for EntitiesConceptsAnalysis.EntitiesConceptsAnalysis:


OA_entities_names.OA _entities_names
EntitiesConceptsAnalysis. EntitiesConceptsAnalysis
EntitiesConceptsAnalysis. AuthorsConceptsAnalysis
EntitiesConceptsAnalysis. ConceptsConceptsAnalysis
EntitiesConceptsAnalysis. InstitutionsConceptsAnalysis
EntitiesConceptsAnalysis. PublishersConceptsAnalysis
EntitiesConceptsAnalysis. SourcesConceptsAnalysis
EntitiesConceptsAnalysis. WorksConceptsAnalysis
EntitiesConceptsPlot.Authors ConceptsPlot
EntitiesConceptsPlot.Concepts ConceptsPlot
EntitiesConceptsPlot.Institutions ConceptsPlot
EntitiesConceptsPlot.Publishers ConceptsPlot
EntitiesConceptsPlot.Sources ConceptsPlot
EntitiesConceptsPlot.Works ConceptsPlot

Collaboration diagram for EntitiesConceptsAnalysis.EntitiesConceptsAnalysis:


OA_entities_names.OA _entities_names
EntitiesConceptsAnalysis. EntitiesConceptsAnalysis

Public Member Functions

- def __init__ (self, entitie_from_id=None, extra_filters=None, database_file_path=None, allow_automatic↩ _download=True, disable_tqdm_loading_bar=False, progress_fct_update=None, create_dataframe=True, entitie_name=None, load_only_columns=None, n_max_entities=default_n_max_entities_to_download)
- def get_count_entities_matched (self, query_filters)

  Gets and return the number of entities which match the query fitlers.
- def get_api_query (self)

  Gets the api query from the parameters of the instance.
- def download_list_entities (self)

  Downloads the entities which match the parameters of the instance, and store the dataset as a parquet file.
- def load_entities_dataframe (self)

  Loads an entities dataset from file (or download it if needed and allowed by the instance) to the dataframe of the instance.
- def get_df_filtered_entities_selection_threshold (self, df_filters)

  Gets df_filtered which contains the entities of self.entities_df fitting the filters in df_filters.
- def get_number_of_entities_selected (self, x_threshold, y_threshold, cited_by_threshold, x_datas, y_datas)

  Gets the number of entities selected on the plot.
- def create_multi_concept_filters_entities_dataframe (self, concepts_from, concepts_filters, thresholds, x_↩ datas, x_threshold, cited_by_threshold)

  Creates the multi concept filters entities dataframe.
- def add_average_combined_concept_score_to_multi_concept_entitie_df (self, concepts_from)

  Adds a column with the average combined concept score to the multi concept entities dataframe.
- def get_database_file_name (self, entitie_from_id=None, entities_type=None, db_format="parquet", extra↩ _text=None)

  Gets the database file name according to the parameters of the ojbect or the arguments given.
- def get_entitie_string_name (self, entitie=None)

  Gets the entitie type string name.
- def get_entitie_type_from_id (self, entitie=None)

  Gets the entitie type from the entitie id.
- def get_name_of_entitie (self, entitie=None, allow_download_from_API=True)

  Gets the name of entitie.

Public Attributes

- **per_page**
- **project_datas_folder_path**
- **entitie_from_id**
- **entitie_from_type**
- **extra_filters**
- **database_file_path**
- **allow_automatic_download**
- **entitie_name**
- **load_only_columns**
- **entities_df**
- **entities_multi_filtered_df**
- **element_count_df**
- **entitie_downloading_progress_percentage**
- **create_element_count_array_progress_percentage**
- **create_element_count_array_progress_text**
- **count_element_type**
- **count_element_years**
- **count_entities_cols**
- **n_max_entities**
- **disable_tqdm_loading_bar**
- **progress_fct_update**
- **cc**

Additional Inherited Members

4.5.1 Detailed Description

OpenAlexAnalysis class which contains generic methods to do analysis over OpenAlex entities.

4.5.2 Constructor & Destructor Documentation

4.5.2.1 __init__()

```
def EntitiesConceptsAnalysis.EntitiesConceptsAnalysis.__init__ (
              self,
              entitie_from_id = None,
              extra_filters = None,
              database_file_path = None,
              allow_automatic_download = True,
              disable_tqdm_loading_bar = False,
              progress_fct_update = None,
              create_dataframe = True,
              entitie_name = None,
              load_only_columns = None,
              n_max_entities = default_n_max_entities_to_download )
```

Parameters

entitie_from_id	The entitie identifier (eg an institution id) from which to take the entities (eg the works) to analyse (str)
filters	Optional additionnal filters, refer to the documentation of openalex and pyalex for the format (dict)
database_file_path	The database file path to force the analyse over datas in a specific file (str)
allow_automatic_download	The allow automatic download (True/False)
disable_tqdm_loading_bar	The disable tqdm loading bar (True/False)
progress_fct_update	The progress fct update UNSUED ???
create_dataframe	Create the dataframe at the initialisation (and download the data if needed and allowed)
entitie_name	To specify the name of the entitie to avoid downloading it via the API if needed

4.5.3 Member Function Documentation

4.5.3.1 add_average_combined_concept_score_to_multi_concept_entitie_df()

```
def EntitiesConceptsAnalysis.EntitiesConceptsAnalysis.add_average_combined_concept_score_to_↩
multi_concept_entitie_df (
              self,
              concepts_from )
```

Adds a column with the average combined concept score to the multi concept entities dataframe.

Parameters

concepts_from	The concepts to use to calculate the combined concept score (list of str)

4.5.3.2 create_multi_concept_filters_entities_dataframe()

```
def EntitiesConceptsAnalysis.EntitiesConceptsAnalysis.create_multi_concept_filters_entities_↩
dataframe (
              self,
              concepts_from,
              concepts_filters,
              thresholds,
              x_datas,
              x_threshold,
              cited_by_threshold )
```

Creates the multi concept filters entities dataframe.

Combines different datasets and filters them.

Parameters

concepts_from	The concept datasets to import and on which the filters will be applied (list of str)
concepts_filters	The concepts which will be used to filter (list of str)
thresholds	The thresholds attached to each concepts to filter (list of float or int)
x_datas	The dataframe key of the global filter (eg the number of works) (str)
x_threshold	The threshold for the global filter (float or int)
cited_by_threshold	The cited by threshold (another global filter) (float or int)

4.5.3.3 get_api_query()

```
def EntitiesConceptsAnalysis.EntitiesConceptsAnalysis.get_api_query (
              self )
```

Gets the api query from the parameters of the instance.

Returns

The api query (dict)

4.5.3.4 get_count_entities_matched()

```
def EntitiesConceptsAnalysis.EntitiesConceptsAnalysis.get_count_entities_matched (
              self,
              query_filters )
```

Gets and return the number of entities which match the query fitlers.

Parameters

query_filters	The query filters (dict)

Returns

The count entities matched (int)

4.5.3.5 get_database_file_name()

```
def EntitiesConceptsAnalysis.EntitiesConceptsAnalysis.get_database_file_name (
              self,
              entitie_from_id = None,
              entities_type = None,
              db_format = "parquet",
              extra_text = None )
```

Gets the database file name according to the parameters of the ojbect or the arguments given.

Parameters

entitie_from↩ _id	The identifier of the entitie (eg a concept id) which was used to filter the entities (eg works) in the database (str)
entities_type	The entities type in the database (eg works) (EntitieOpenAlex)
db_format	The database file format (str)
extra_text	Extra text to add to the file name (str)

Returns

The database file name (str)

4.5.3.6 get_df_filtered_entities_selection_threshold()

```
def EntitiesConceptsAnalysis.EntitiesConceptsAnalysis.get_df_filtered_entities_selection_↩
threshold (
              self,
              df_filters )
```

Gets df_filtered which contains the entities of self.entities_df fitting the filters in df_filters.

Parameters

df_filters	The filters in a dictionnary with for the key for the data to filter and for the value the minimum threshold (dict)

Returns

df_filtered, corresponding the the entities fitting the thresholds (pandas DataFrame)

4.5.3.7 get_entitie_string_name()

```
def EntitiesConceptsAnalysis.EntitiesConceptsAnalysis.get_entitie_string_name (
              self,
              entitie = None )
```

Gets the entitie type string name.

Parameters

entitie	The entitie, if not provided, the instance entitie id will be used (BaseOpenAlex)

Returns

The entitie type name (str)

4.5.3.8 get_entitie_type_from_id()

```
def EntitiesConceptsAnalysis.EntitiesConceptsAnalysis.get_entitie_type_from_id (
              self,
              entitie = None )
```

Gets the entitie type from the entitie id.

Parameters

entitie	The entitie id (str)

Returns

The entitie type (BaseOpenAlex)

4.5.3.9 get_name_of_entitie()

```
def EntitiesConceptsAnalysis.EntitiesConceptsAnalysis.get_name_of_entitie (
              self,
              entitie = None,
              allow_download_from_API = True )
```

Gets the name of entitie.

Parameters

entitie	The entitie id, if not provided, use the one from the instance (str)
allow_download_from_API	Allow to download the entitie name from the OpenAlex API (bool)

Returns

The name of entitie (str)

4.5.3.10 get_number_of_entities_selected()

```
def EntitiesConceptsAnalysis.EntitiesConceptsAnalysis.get_number_of_entities_selected (
              self,
              x_threshold,
              y_threshold,
              cited_by_threshold,
              x_datas,
              y_datas )
```

Gets the number of entities selected on the plot.

Parameters

x_threshold	The x threshold (float or int)
y_threshold	The y threshold (float or int)
cited_by_threshold	The cited by threshold (float or int)
x_datas	The x datas key on the dataframe (str)
y_datas	The y datas key on the dataframe (str)

Returns

The number of entities selected (int)

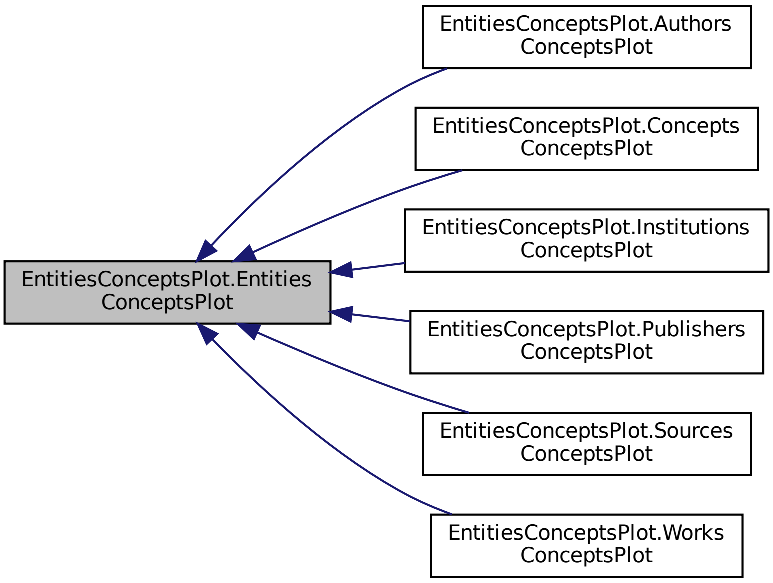
The documentation for this class was generated from the following file:

- EntitiesConceptsAnalysis.py

4.6 EntitiesConceptsPlot.EntitiesConceptsPlot Class Reference

EntitiesConceptsPlot class which contains generic methods to do plots of OpenAlex entities.

Inheritance diagram for EntitiesConceptsPlot.EntitiesConceptsPlot:


EntitiesConceptsPlot.Authors ConceptsPlot
EntitiesConceptsPlot.Concepts ConceptsPlot
EntitiesConceptsPlot.Institutions ConceptsPlot
EntitiesConceptsPlot.Entities ConceptsPlot
EntitiesConceptsPlot.Publishers ConceptsPlot
EntitiesConceptsPlot.Sources ConceptsPlot
EntitiesConceptsPlot.Works ConceptsPlot

Public Member Functions

- def get_figure_entities_of_a_concept_color_country (self, concept, plot_parameters=None)

  Gets the figure with the entities of a concept, and with the country as color.
- def get_figure_entities_selection_threshold (self, concept, plot_parameters, x_threshold=0, y_threshold=0, cited_by_threshold=0, display_only_selected_entities=None, display_threshold_lines=None, entity_to_↩ highlight=None)

  Gets the figure with the entities of a concept and the selection threshold lines (optional)
- def get_figure_time_series_element_used_by_entities (self, element=None, plot_title=None, x_datas='year', x_legend="Year", y_datas=None, color_legend="Entities")

  Gets the figure with the time series usage of a element (eg reference, concept) by entities.

4.6.1 Detailed Description

EntitiesConceptsPlot class which contains generic methods to do plots of OpenAlex entities.

4.6.2 Member Function Documentation

4.6.2.1 get_figure_entities_of_a_concept_color_country()

```
def EntitiesConceptsPlot.EntitiesConceptsPlot.get_figure_entities_of_a_concept_color_country (
              self,
              concept,
              plot_parameters = None )
```

Gets the figure with the entities of a concept, and with the country as color.

Parameters

concept	The concept (str)
plot_parameters	The plot parameters (dict)

Returns

The figure (fig)

4.6.2.2 get_figure_entities_selection_threshold()

```
def EntitiesConceptsPlot.EntitiesConceptsPlot.get_figure_entities_selection_threshold (
              self,
              concept,
              plot_parameters,
              x_threshold = 0,
              y_threshold = 0,
              cited_by_threshold = 0,
              display_only_selected_entities = None,
              display_threshold_lines = None,
              entity_to_highlight = None )
```

Gets the figure with the entities of a concept and the selection threshold lines (optional)

Parameters

concept	The concept (str)
plot_parameters	The plot parameters (dict)
x_threshold	The x threshold (float or int)
y_threshold	The y threshold (float or int)
cited_by_threshold	The cited by threshold (float or int)
display_only_selected_entities	The display only selected entities (bool)
display_threshold_lines	The display threshold lines (bool)
entity_to_highlight	The entity to highlight on the plot (str)

Returns

The figure (fig)

4.6.2.3 get_figure_time_series_element_used_by_entities()

```
def EntitiesConceptsPlot.EntitiesConceptsPlot.get_figure_time_series_element_used_by_entities
(
              self,
              element = None,
              plot_title = None,
              x_datas = 'year',
              x_legend = "Year",
              y_datas = None,
              color_legend = "Entities" )
```

Gets the figure with the time series usage of a element (eg reference, concept) by entities.

Parameters

element	The element (default first in the dataframe) (str)
plot_title	The plot title (str)
x_datas	The x datas (default: year) (str)
x_legend	The x legend (str)
y_datas	The y datas (the entities to plot, the default is all entities in the dataframe) (list[str])
color_legend	The color legend (str)

Returns

The figure (fig)

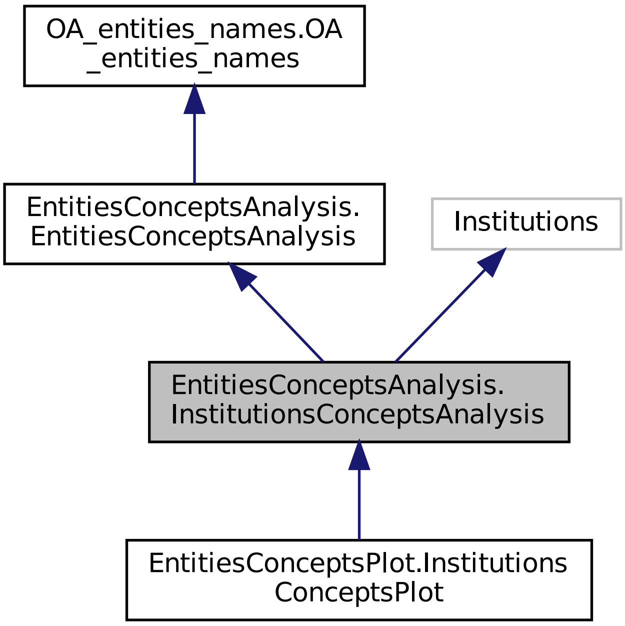
The documentation for this class was generated from the following file:

- EntitiesConceptsPlot.py

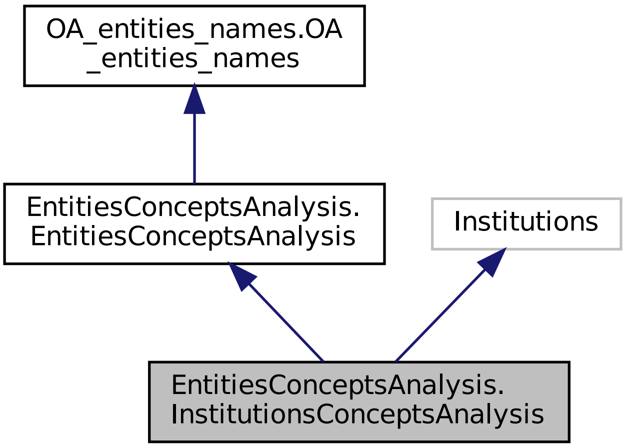
4.7 EntitiesConceptsAnalysis.InstitutionsConceptsAnalysis Class Reference

This class contains specific methods for Institutions concepts analysis.

Inheritance diagram for EntitiesConceptsAnalysis.InstitutionsConceptsAnalysis:


OA_entities_names.OA
_entities_names
EntitiesConceptsAnalysis.
EntitiesConceptsAnalysis
Institutions
EntitiesConceptsAnalysis.
InstitutionsConceptsAnalysis
EntitiesConceptsPlot.Institutions
ConceptsPlot

Collaboration diagram for EntitiesConceptsAnalysis.InstitutionsConceptsAnalysis:


OA_entities_names.OA
_entities_names
EntitiesConceptsAnalysis.
EntitiesConceptsAnalysis
Institutions
EntitiesConceptsAnalysis.
InstitutionsConceptsAnalysis

Public Member Functions

- def filter_and_format_entitie_data_from_api_response (self, entitie)

Filter and format the institutions data downloaded from the API.

- def get_sum_concept_scores (self, institutions, concept_links)

  Gets the sum of the concept scores of the concepts in the list concepts.

Public Attributes

- **entitie_from_type**

Static Public Attributes

- **EntitieOpenAlex** = Institutions

4.7.1 Detailed Description

This class contains specific methods for Institutions concepts analysis.

4.7.2 Member Function Documentation

4.7.2.1 filter_and_format_entitie_data_from_api_response()

```
def EntitiesConceptsAnalysis.InstitutionsConceptsAnalysis.filter_and_format_entitie_data_↩
from_api_response (
              self,
              entitie )
```

Filter and format the institutions data downloaded from the API.

Parameters

entitie	The institutions data from the API (dict)

Returns

The institutions datas (dict)

4.7.2.2 get_sum_concept_scores()

```
def EntitiesConceptsAnalysis.InstitutionsConceptsAnalysis.get_sum_concept_scores (
              self,
              institutions,
              concept_links )
```

Gets the sum of the concept scores of the concepts in the list concepts.

Parameters

Institutions	The institution (list of dict)
concept_links	The concept links

Returns

The sum of the concept scores

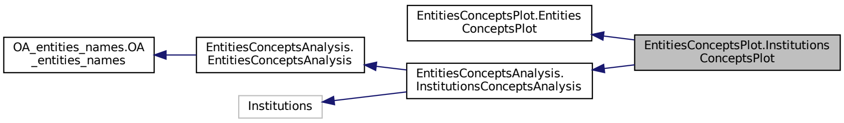
The documentation for this class was generated from the following file:

- EntitiesConceptsAnalysis.py

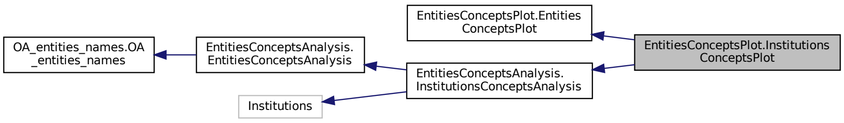
4.8 EntitiesConceptsPlot.InstitutionsConceptsPlot Class Reference

This class contains specific methods for Institutions concepts plot.

Inheritance diagram for EntitiesConceptsPlot.InstitutionsConceptsPlot:


EntitiesConceptsPlot.Entities ConceptsPlot
OA_entities_names.OA _entities_names
EntitiesConceptsAnalysis. EntitiesConceptsAnalysis
EntitiesConceptsPlot.Institutions ConceptsPlot
EntitiesConceptsAnalysis. InstitutionsConceptsAnalysis
Institutions

Collaboration diagram for EntitiesConceptsPlot.InstitutionsConceptsPlot:


EntitiesConceptsPlot.Entities ConceptsPlot
OA_entities_names.OA _entities_names
EntitiesConceptsAnalysis. EntitiesConceptsAnalysis
EntitiesConceptsPlot.Institutions ConceptsPlot
EntitiesConceptsAnalysis. InstitutionsConceptsAnalysis
Institutions

Public Member Functions

- def getCustomData (self, concept)

  Gets the custom data for the plot.
- def getHoverTemplate (self, concept)

  Gets the hover template for the plot.
- def get_figure_institutions_multi_concepts_filtered (self, plot_parameters, concepts_from, concepts_filters, thresholds, x_threshold, cited_by_threshold, institution_to_highlight)

  Gets the figure with the institutions of multiple concepts and filtered.

Additional Inherited Members

4.8.1 Detailed Description

This class contains specific methods for Institutions concepts plot.

4.8.2 Member Function Documentation

4.8.2.1 get_figure_institutions_multi_concepts_filtered()

```
def EntitiesConceptsPlot.InstitutionsConceptsPlot.get_figure_institutions_multi_concepts_←
filtered (
              self,
              plot_parameters,
              concepts_from,
              concepts_filters,
              thresholds,
              x_threshold,
              cited_by_threshold,
              institution_to_highlight )
```

Gets the figure with the institutions of multiple concepts and filtered.

Parameters

plot_parameters	The plot parameters (dict)
concepts_from	The concepts to import to create the dataset (list of str)
concepts_filters	The concepts to use to filter the institutions (list of str)
thresholds	The thresholds for each concept filter (list of float or int)
x_threshold	The global threshold (eg nb of works), usually corresponding to the x data (float or int)
cited_by_threshold	The cited by threshold (float or int)
institution_to_highlight	The institution to highlight on the plot (str)

Returns

The figure (fig)

4.8.2.2 getCustomData()

```
def EntitiesConceptsPlot.InstitutionsConceptsPlot.getCustomData (
              self,
              concept )
```

Gets the custom data for the plot.

Parameters

concept	The concept (str)

Returns

The custom data (list of str)

4.8.2.3 getHoverTemplate()

```
def EntitiesConceptsPlot.InstitutionsConceptsPlot.getHoverTemplate (
              self,
              concept )
```

Gets the hover template for the plot.

Parameters

concept	The concept (str)

Returns

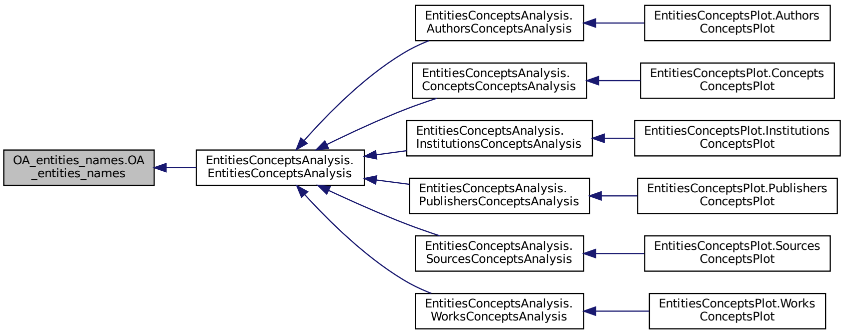
The hover template (list of str)

The documentation for this class was generated from the following file:

- EntitiesConceptsPlot.py

4.9 OA_entities_names.OA_entities_names Class Reference

Inheritance diagram for OA_entities_names.OA_entities_names:


EntitiesConceptsAnalysis.
AuthorsConceptsAnalysis
EntitiesConceptsPlot.Authors
ConceptsPlot
EntitiesConceptsAnalysis.
ConceptsConceptsAnalysis
EntitiesConceptsPlot.Concepts
ConceptsPlot
EntitiesConceptsAnalysis.
InstitutionsConceptsAnalysis
EntitiesConceptsPlot.Institutions
ConceptsPlot
OA_entities_names.OA
_entities_names
EntitiesConceptsAnalysis.
EntitiesConceptsAnalysis
EntitiesConceptsAnalysis.
PublishersConceptsAnalysis
EntitiesConceptsPlot.Publishers
ConceptsPlot
EntitiesConceptsAnalysis.
SourcesConceptsAnalysis
EntitiesConceptsPlot.Sources
ConceptsPlot
EntitiesConceptsAnalysis.
WorksConceptsAnalysis
EntitiesConceptsPlot.Works
ConceptsPlot

Public Member Functions

- def **__init__** (self)
- def **get_concepts_institutions_id_downloaded** (self)
- def get_concepts_institutions_names_downloaded (self)

  Institutions ### create the list of all the concept id with the databases downloaded.
- def **get_concepts_institutions_names_downloadable** (self)

Static Public Attributes

- string concepts_csv_file = "OpenAlex_concepts_in_use_(17_August_2022)_-_concepts.csv"

  CONCEPTS #####.
- string **concepts_institutions_database_files_directory** = "data/"
- int **max_concept_level** = 2
- string **databases_format** = ".parquet"
- **df_concepts** = pd.read_csv(concepts_csv_file)
- **concepts_names_full** = df_concepts[['openalex_id', 'display_name']].set_index('openalex_id')['display_↩ name'].to_dict()
- **concepts_levels_full** = df_concepts[['openalex_id', 'level']].set_index('openalex_id')['level'].to_dict()
- **df_concepts_normalized_names_full** = df_concepts[['openalex_id', 'normalized_name']]
- **regex**
- **concepts_names** = df_concepts[['openalex_id', 'display_name']].set_index('openalex_id')['display_↩ name'].to_dict()
- **df_concepts_normalized_names** = df_concepts[['openalex_id', 'normalized_name']]
- **df_concepts_institutions_file_names** = df_concepts[['openalex_id', 'normalized_name_institutions']]
- **concepts_normalized_names** = df_concepts_normalized_names.set_index('openalex_id')['normalized_↩ name'].to_dict()
- **concepts_institutions_database_file_name** = df_concepts_institutions_file_names.set_index('openalex↩ _id')['normalized_name_institutions'].to_dict()
- string list_of_institutions_file_path = "list_all_institutions.parquet"

  INSTITUTIONS #####.
- **institutions_df** = pd.read_parquet(list_of_institutions_file_path, columns = ['id', 'display_name'])
- **institutions_names** = institutions_df[['id', 'display_name']].set_index('id')['display_name'].to_dict()

4.9.1 Detailed Description

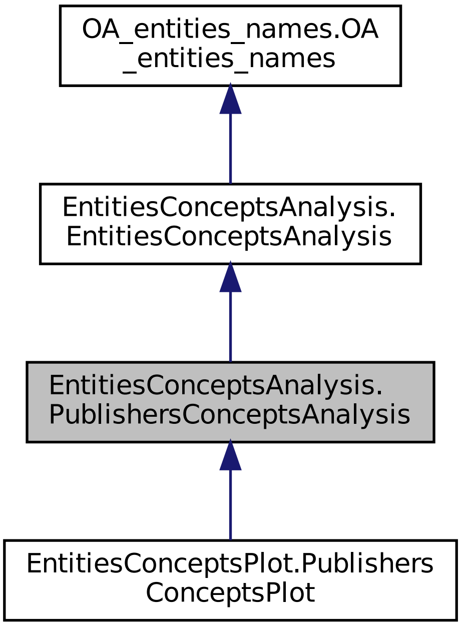
```
Class to manage the concept names and databases file names for the webapp
```

The documentation for this class was generated from the following file:

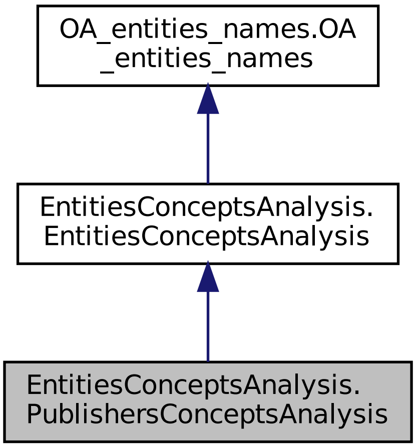
- OA_entities_names.py

4.10 EntitiesConceptsAnalysis.PublishersConceptsAnalysis Class Reference

Inheritance diagram for EntitiesConceptsAnalysis.PublishersConceptsAnalysis:


OA_entities_names.OA _entities_names
EntitiesConceptsAnalysis. EntitiesConceptsAnalysis
EntitiesConceptsAnalysis. PublishersConceptsAnalysis
EntitiesConceptsPlot.Publishers ConceptsPlot

Collaboration diagram for EntitiesConceptsAnalysis.PublishersConceptsAnalysis:


OA_entities_names.OA _entities_names
EntitiesConceptsAnalysis. EntitiesConceptsAnalysis
EntitiesConceptsAnalysis. PublishersConceptsAnalysis

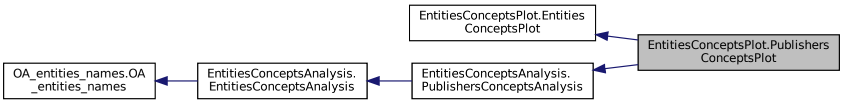
Additional Inherited Members

The documentation for this class was generated from the following file:

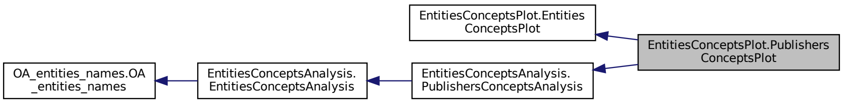
- EntitiesConceptsAnalysis.py

4.11 EntitiesConceptsPlot.PublishersConceptsPlot Class Reference

Inheritance diagram for EntitiesConceptsPlot.PublishersConceptsPlot:


EntitiesConceptsPlot.Entities ConceptsPlot
EntitiesConceptsPlot.Publishers ConceptsPlot
OA_entities_names.OA _entities_names
EntitiesConceptsAnalysis. EntitiesConceptsAnalysis
EntitiesConceptsAnalysis. PublishersConceptsAnalysis

Collaboration diagram for EntitiesConceptsPlot.PublishersConceptsPlot:


EntitiesConceptsPlot.Entities ConceptsPlot
EntitiesConceptsPlot.Publishers ConceptsPlot
OA_entities_names.OA _entities_names
EntitiesConceptsAnalysis. EntitiesConceptsAnalysis
EntitiesConceptsAnalysis. PublishersConceptsAnalysis

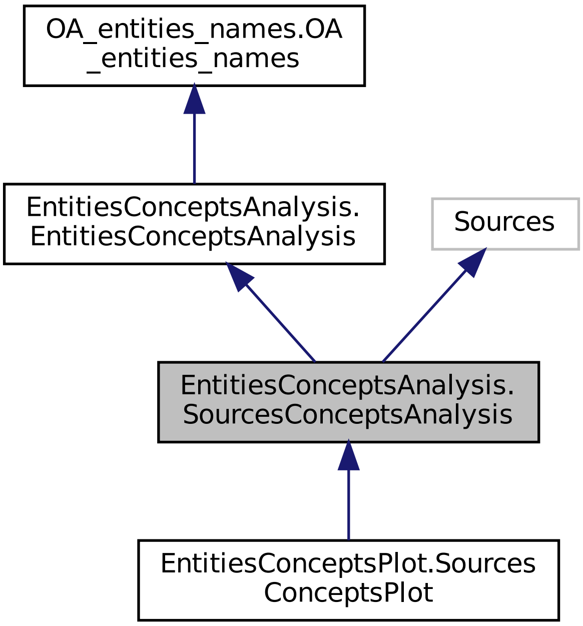
Additional Inherited Members

The documentation for this class was generated from the following file:

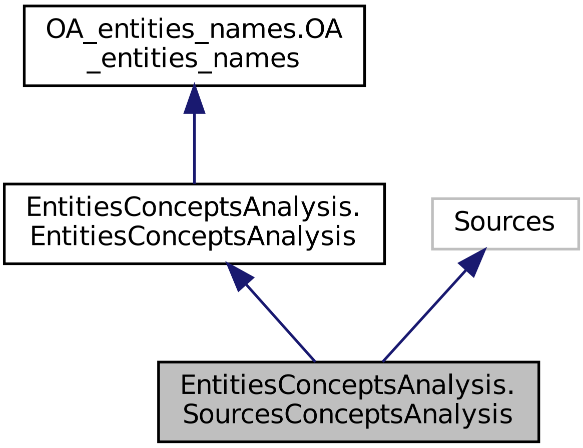
- EntitiesConceptsPlot.py

4.12 EntitiesConceptsAnalysis.SourcesConceptsAnalysis Class Reference

Inheritance diagram for EntitiesConceptsAnalysis.SourcesConceptsAnalysis:



Collaboration diagram for EntitiesConceptsAnalysis.SourcesConceptsAnalysis:



Static Public Attributes

- **EntitieOpenAlex** = Sources

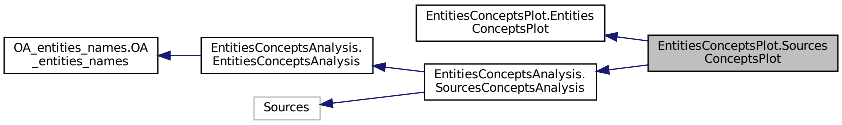
Additional Inherited Members

The documentation for this class was generated from the following file:

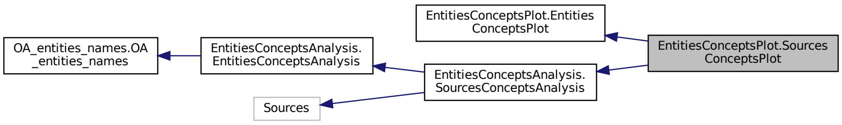
- EntitiesConceptsAnalysis.py

4.13 EntitiesConceptsPlot.SourcesConceptsPlot Class Reference

Inheritance diagram for EntitiesConceptsPlot.SourcesConceptsPlot:


OA_entities_names.OA _entities_names
EntitiesConceptsAnalysis. EntitiesConceptsAnalysis
Sources
EntitiesConceptsPlot.Entities ConceptsPlot
EntitiesConceptsAnalysis. SourcesConceptsAnalysis
EntitiesConceptsPlot.Sources ConceptsPlot

Collaboration diagram for EntitiesConceptsPlot.SourcesConceptsPlot:


OA_entities_names.OA _entities_names
EntitiesConceptsAnalysis. EntitiesConceptsAnalysis
Sources
EntitiesConceptsPlot.Entities ConceptsPlot
EntitiesConceptsAnalysis. SourcesConceptsAnalysis
EntitiesConceptsPlot.Sources ConceptsPlot

Additional Inherited Members

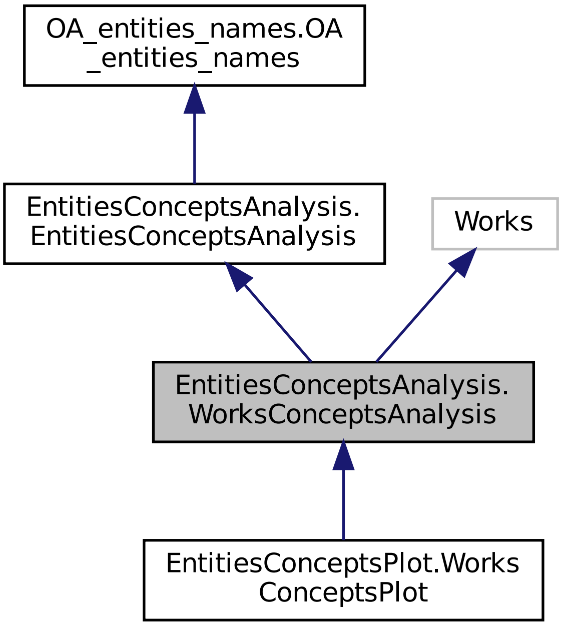
The documentation for this class was generated from the following file:

- EntitiesConceptsPlot.py

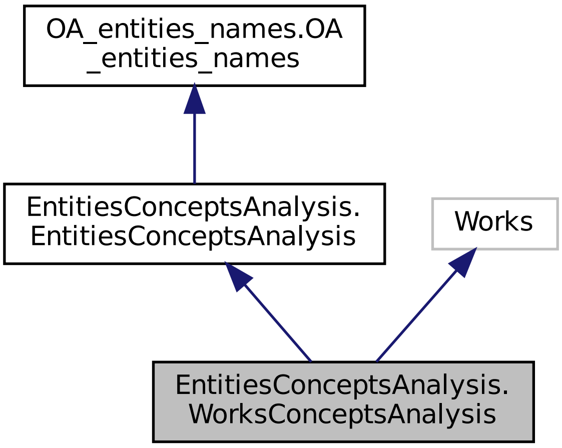
4.14 EntitiesConceptsAnalysis.WorksConceptsAnalysis Class Reference

This class contains specific methods for Works concepts analysis.

Inheritance diagram for EntitiesConceptsAnalysis.WorksConceptsAnalysis:



Collaboration diagram for EntitiesConceptsAnalysis.WorksConceptsAnalysis:



Public Member Functions

- def filter_and_format_entitie_data_from_api_response (self, entitie)

  Filter and format the works data downloaded from the API.
- def get_country_code (self, entitie)

  Gets the country code from an entitie.
- def get_institution_name (self, entitie)

  Gets the institution name from an entitie.
- def get_works_references_count (self, count_years=[ ])

  Gets the works references count of the works list of the instance.
- def get_works_concepts_count (self, count_years=[ ])

  Gets the concepts count of the works list of the instance.
- def **get_element_count** (self, element_type, count_years=[ ])
- def create_element_used_count_array (self, element_type, entities_from=[ ], out_file_name=None, save_↩ out_array=False, count_years=[ ])

  Creates the element used count array.
- def sort_count_array (self, sort_by='h_used_all_l_use_main', sort_by_ascending=False)

  Sort the dataframe with the count array (element_count_df)
- def add_statistics_to_element_count_array (self, sort_by='h_used_all_l_use_main', sort_by_ascending=False, min_concept_level=None)

  Adds a statistics to the element count array (statistics between the main entitie to compare (second column in the dataframe) and the sum of the other entities)
- def add_statistics_to_references_works_count_array (self)

  Adds a statistics to the references works count array (statistics between the main entitie to compare (second column in the dataframe) and the sum of the other entities)
- def add_statistics_to_concept_count_array (self, min_concept_level=None)

  Adds a statistics to the concepts count array (statistics between the main entitie to compare (second column in the dataframe) and the sum of the other entities)

Public Attributes

- **entitie_from_type**
- **count_element_type**
- **count_element_years**
- **count_entities_cols**
- **entitie_from_id**
- **element_count_df**
- **create_element_count_array_progress_percentage**
- **create_element_count_array_progress_text**

Static Public Attributes

- **EntitieOpenAlex** = Works

4.14.1 Detailed Description

This class contains specific methods for Works concepts analysis.

4.14.2 Member Function Documentation

4.14.2.1 add_statistics_to_element_count_array()

```
def EntitiesConceptsAnalysis.WorksConceptsAnalysis.add_statistics_to_element_count_array (
              self,
              sort_by = 'h_used_all_l_use_main',
              sort_by_ascending = False,
              min_concept_level = None )
```

Adds a statistics to the element count array (statistics between the main entitie to compare (second column in the dataframe) and the sum of the other entities)

Parameters

sort_by	The key to sort the dataframe (str)
sort_by_ascending	Whenever to sort the dataframe ascending (bool)
min_concept_level	In case the element is a concept, this is the minimum level of the concepts we will keep (aka remove the lower (= more global) concepts)

4.14.2.2 create_element_used_count_array()

```
def EntitiesConceptsAnalysis.WorksConceptsAnalysis.create_element_used_count_array (
              self,
              element_type,
              entities_from = [],
              out_file_name = None,
              save_out_array = False,
              count_years = [] )
```

Creates the element used count array.

Count the number of times each element (eg reference, concept..) is used and save the array as CSV (optional)

Parameters

element_type	The element type
entities_from	The extra entities to which to count the concepts (list of str)
out_file_name	The out CSV file name, if not provided, an appropriate name is generated (str)
save_out_array	Save out array (bool)
count_years	If given, it will compute the count for each year (list[int])

4.14.2.3 filter_and_format_entitie_data_from_api_response()

```
def EntitiesConceptsAnalysis.WorksConceptsAnalysis.filter_and_format_entitie_data_from_api_↩
response (
              self,
              entitie )
```

Filter and format the works data downloaded from the API.

Parameters

entitie	The works data from the API (dict)

Returns

The works datas (dict)

4.14.2.4 get_country_code()

```
def EntitiesConceptsAnalysis.WorksConceptsAnalysis.get_country_code (
              self,
              entitie )
```

Gets the country code from an entitie.

Parameters

entitie	The entitie (dict)

Returns

The country code (str)

4.14.2.5 get_institution_name()

```
def EntitiesConceptsAnalysis.WorksConceptsAnalysis.get_institution_name (
              self,
              entitie )
```

Gets the institution name from an entitie.

Parameters

entitie	The entitie (dict)

Returns

The institution name (str)

4.14.2.6 get_works_concepts_count()

```
def EntitiesConceptsAnalysis.WorksConceptsAnalysis.get_works_concepts_count (
              self,
              count_years = [] )
```

Gets the concepts count of the works list of the instance.

Parameters

count_only_year	If different than None, count only the concepts of the works of the given year (int)

Returns

The concept count (pandas Serie)

4.14.2.7 get_works_references_count()

```
def EntitiesConceptsAnalysis.WorksConceptsAnalysis.get_works_references_count (
              self,
              count_years = [] )
```

Gets the works references count of the works list of the instance.

Returns

The works references count (pandas Serie)

4.14.2.8 sort_count_array()

```
def EntitiesConceptsAnalysis.WorksConceptsAnalysis.sort_count_array (
              self,
              sort_by = 'h_used_all_l_use_main',
              sort_by_ascending = False )
```

Sort the dataframe with the count array (element_count_df)

Parameters

sort_by	The key to sort the dataframe (str)
sort_by_ascending	Whenever to sort the dataframe ascending (bool)

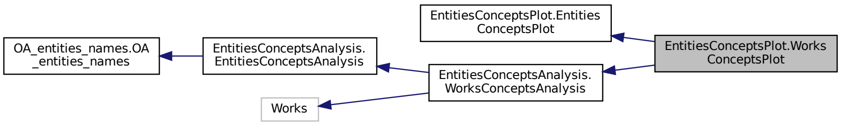
The documentation for this class was generated from the following file:

- EntitiesConceptsAnalysis.py

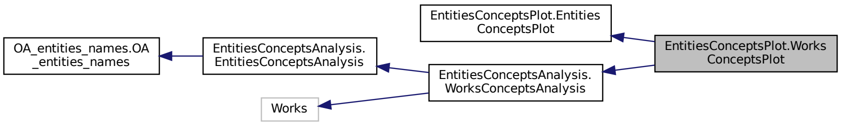
4.15 EntitiesConceptsPlot.WorksConceptsPlot Class Reference

This class contains specific methods for Works concepts plot.

Inheritance diagram for EntitiesConceptsPlot.WorksConceptsPlot:


EntitiesConceptsPlot.Entities ConceptsPlot
EntitiesConceptsPlot.Works ConceptsPlot
OA_entities_names.OA _entities_names
EntitiesConceptsAnalysis. EntitiesConceptsAnalysis
EntitiesConceptsAnalysis. WorksConceptsAnalysis
Works

Collaboration diagram for EntitiesConceptsPlot.WorksConceptsPlot:


EntitiesConceptsPlot.Entities ConceptsPlot
EntitiesConceptsPlot.Works ConceptsPlot
OA_entities_names.OA _entities_names
EntitiesConceptsAnalysis. EntitiesConceptsAnalysis
EntitiesConceptsAnalysis. WorksConceptsAnalysis
Works

Public Member Functions

- def getCustomData (self, concept)

  Gets the custom data for the plot.
- def getHoverTemplate (self, concept)

  Gets the hover template for the plot.
- def get_figure_nb_time_referenced (self, element_type)

  Gets the figure with the number of time each reference is used in a list of works.

Additional Inherited Members

4.15.1 Detailed Description

This class contains specific methods for Works concepts plot.

4.15.2 Member Function Documentation

4.15.2.1 get_figure_nb_time_referenced()

```
def EntitiesConceptsPlot.WorksConceptsPlot.get_figure_nb_time_referenced (
              self,
              element_type )
```

Gets the figure with the number of time each reference is used in a list of works.

Parameters

element_type	The element type

Returns

The figure works number of time referenced (fig)

4.15.2.2 getCustomData()

```
def EntitiesConceptsPlot.WorksConceptsPlot.getCustomData (
              self,
              concept )
```

Gets the custom data for the plot.

Parameters

concept	The concept (str)

Returns

The custom data (list of str)

4.15.2.3 getHoverTemplate()

```
def EntitiesConceptsPlot.WorksConceptsPlot.getHoverTemplate (
              self,
              concept )
```

Gets the hover template for the plot.

Parameters

concept	The concept (str)

Returns

The hover template (list of str)

The documentation for this class was generated from the following file:

- EntitiesConceptsPlot.py