



Класс

Предок

Модуль

Назнач

Chemb1DocumentPipeline

Chemb1CommonPipeline

`src/bioetl/pipelines/chembl/
document/run.py`

Скелет
Документ
источни
специа
валида
для дан

Класс	Предок	Модуль	Назначение
-------	--------	--------	------------

ChembICommonPipeline

ChembIPipelineBase

src/bioetl/pipelines/chembl/
common/base.py

Базовый
пайплайн
10 . От
проверки
инициа
результ
извлече

Класс	Предок	Модуль	Назначение
-------	--------	--------	------------

ChembIPipelineBase

PipelineBase

src/bioetl/core/pipeline/
unified.py

Базовый класс
делегирует
сервису
извлечения
определение
задаёт сущности
для первого

Класс	Предок	Модуль	Назначение
PipelineBase	PipelineRuntimeBase	src/bioetl/core/pipeline/unified.py	<p>Абстрактный класс для определения пайплайна.</p> <p>Этапы ETL:</p> <ul style="list-style-type: none"> save_result умолчательные значения <p>Panderization:</p> <ul style="list-style-type: none"> результаты сервисы
DocumentSchema	pa.DataFrameSchema (Pandera)	src/bioetl/core/schemas/document_schema.py	<p>Схема для валидации DataFrame.</p> <p>Обязательные поля:</p> <ul style="list-style-type: none"> (document_id) journal_id <p>Использование:</p> <ul style="list-style-type: none"> набора

Класс	Предок	Модуль	Назначение
ChEMBLExtractionService	(нет)	<code>src/bioetl/pipelines/chembl/common/chembl_extraction_service.py</code>	Сервис, шаги из 51 . Упаковка, формирование, побатчинг, клиент, объединение, результат
ChEMBLExtractionServiceDescriptor	(нет)	<code>src/bioetl/core/pipeline/unified.py</code>	Описано сущность 69 . Со, построение (build), (fetch), (final), ChEMBL получение

Класс	Предок	Модуль	Назначение
Chemb1DescriptorFactory	(нет)	<code>src/bioetl/clients/chembl/ descriptor_factory.py</code>	Фабрика основе (Chembl) выборка дескриптор сущности пагинация выборка
ExtractionStrategyFactory	(нет)	<code>src/bioetl/pipelines/chembl/ commonestrategies.py</code>	Фабрика данных реализации dataclasses дескриптор подхода от типа конфигурации
ServiceExtractionStrategy	(нет)	<code>src/bioetl/pipelines/chembl/ commonestrategies.py</code>	Стратегия дескриптор выборка типа "сервис" пайплайн (если нет run_de получен возвраща результат если выбы

Класс	Предок	Модуль	Назначение
ChEMBLWriteService	(нет)	<code>src/bioetl/pipelines/chembl/common/base.py</code>	Детерминирует ChEMBL утилиты для сохранения файлов с именем (<entity>.формат) и YAML обеспечивая единство
DefaultValidationService	(нет)	<code>src/bioetl/core/pipeline/services/defaults.py</code>	Стандартные DataFrame схемы DataFrame схеме (если сортируя результат)
PipelineOutputService	(нет)	<code>src/bioetl/core/io/output_service.py</code>	Сервис пайплайнов. Пытается использовать writer (работа с данными) (например, в БД); если генерируется об отсутствии (что заставляет использовать записи)

Класс	Предок	Модуль	Назначение
ConfigValidationError	<code>ValueError</code>	<code>src/bioetl/pipelines/chembl/common/descriptor.py</code>	Исключение некорректного конфигурационного файла. Используется для неверных значений, обязательных (например, т.д.), пропущенных или пайплайнами.

1 2 3 4 5 6 7 8 9 **run.py**

https://github.com/SatoryKono/bioactivity_data_acquisition/blob/98f17f432532cddd56d0a07fd4796b37c54677ec/src/bioetl/pipelines/chembl/document/run.py

10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 79 80 81 82 83 **base.py**

https://github.com/SatoryKono/bioactivity_data_acquisition/blob/98f17f432532cddd56d0a07fd4796b37c54677ec/src/bioetl/pipelines/chembl/common/base.py

27 28 71 72 73 74 75 76 77 78 **strategies.py**

https://github.com/SatoryKono/bioactivity_data_acquisition/blob/98f17f432532cddd56d0a07fd4796b37c54677ec/src/bioetl/pipelines/chembl/common/strategies.py

29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 69 **unified.py**

https://github.com/SatoryKono/bioactivity_data_acquisition/blob/98f17f432532cddd56d0a07fd4796b37c54677ec/src/bioetl/core/pipeline/unified.py

50 **document_schema.py**

https://github.com/SatoryKono/bioactivity_data_acquisition/blob/98f17f432532cddd56d0a07fd4796b37c54677ec/src/bioetl/core/schemas/document_schema.py

51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 **chembl_extraction_service.py**

https://github.com/SatoryKono/bioactivity_data_acquisition/blob/98f17f432532cddd56d0a07fd4796b37c54677ec/src/bioetl/pipelines/chembl/common/chembl_extraction_service.py

70 **descriptor_factory.py**

https://github.com/SatoryKono/bioactivity_data_acquisition/blob/98f17f432532cddd56d0a07fd4796b37c54677ec/src/bioetl/clients/chembl/descriptor_factory.py

84 85 86 87 **defaults.py**

https://github.com/SatoryKono/bioactivity_data_acquisition/blob/98f17f432532cddd56d0a07fd4796b37c54677ec/src/bioetl/core/pipeline/services/defaults.py

88 89 90 91 92 **output_service.py**

https://github.com/SatoryKono/bioactivity_data_acquisition/blob/98f17f432532cddd56d0a07fd4796b37c54677ec/src/bioetl/core/io/output_service.py

93 **descriptor.py**

https://github.com/SatoryKono/bioactivity_data_acquisition/blob/98f17f432532cddd56d0a07fd4796b37c54677ec/src/bioetl/pipelines/chembl/common/descriptor.py