libera_utils

Release 5.0.0

Libera SDC Team

Nov 07, 2025

CONTENTS:

Version: 5.0.0

CHAPTER

ONE

USER DOCS

This section is for developers including Libera Utils in their code. If you are a Libera L2 algorithm developer, you are in the right place.

1.1 Basic Usage

1.1.1 Command Line Interface

The CLI is installed as an executable in your virtual environment during installation of `libera_utils`.

Top Level Command `libera-utils`

This is the top level command that contains all the nested sub-commands.

```
usage: libera-utils [-h] [--version] {make-kernel,ecr-upload,step-function-trigger} ...

Libera SDC utilities CLI

options:
  -h, --help            show this help message and exit
  --version             print current version of the CLI

subcommands:
  sub-commands for libera-utils CLI

  {make-kernel,ecr-upload,step-function-trigger}
    make-kernel         generate SPICE kernel from telemetry data
    ecr-upload          Upload docker image to ECR repository for a specific algorithm
    step-function-trigger
                        Manually trigger a specific step function
```

Sub-Command `ecr-upload`

This is a tool to upload a docker image to AWS ECR. The image name and tag identify the local docker image while the –ecr-image-tags option specifies the tags to apply to the image in the ECR (remote tags). If `--ecr-image-tags` is not provided, only the `latest` tag is applied by default. If `--ecr-image-tags` is specified, you must include `latest` explicitly.

```
usage: libera-utils ecr-upload [-h] [--ecr-image-tags ECR_IMAGE_TAGS [ECR_IMAGE_TAGS ...
↪]] [--ignore-docker-config] image_name image_tag algorithm_name
```

(continues on next page)

(continued from previous page)

```
positional arguments:
  image_name            Image name of image to upload (image-name:image-tag)
  image_tag             Image tag of image to upload (image-name:image-tag)
  algorithm_name        Algorithm name that matches an ECR repo name, inputs to names:
                        ['cal-rad', 'cal-cam', 'spice-azel', 'spice-jpss', 'l1b-rad',
↪'l1b-cam', 'int-footprint-scene-id',
                        'l2-cf-rad', 'l2-cf-cam', 'l2-unfiltered', 'l2-ssw-toa-osse',
↪'l2-ssw-toa-erbe', 'l2-ssw-toa-trmm',
                        'l2-ssw-toa-rt', 'l2-ssw-surface-flux', 'adm-binning']

options:
  -h, --help            show this help message and exit
  --ecr-image-tags ECR_IMAGE_TAGS [ECR_IMAGE_TAGS ...]
                        List of tags to apply to the uploaded image in the ECR (e.g. `--
↪ecr-image-tags latest 1.3.4`) Note, latest is applied if this option is not set. If it␣
↪is set, you must specify
                        latest if you want it tagged as such in the ECR.
  --ignore-docker-config
                        Ignore the standard docker config.json to bypass the credential␣
↪store
```

Example usage:

```
libera-utils ecr-upload recently-built-ssw-sfc-flux latest l2-ssw-surface-flux --ecr-
↪image-tags latest --ignore-docker-config
```

To get a list of specific algorithm names allowed in this command, run `libera-utils ecr-upload -h`

Sub-Command `make-kernel jpss-spk`

```
usage: libera-utils make-kernel jpss-spk [-h] --outdir OUTDIR [--overwrite] [-v] packet_
↪data_filepaths [packet_data_filepaths ...]

positional arguments:
  packet_data_filepaths
                        paths to L0 packet files

options:
  -h, --help            show this help message and exit
  --outdir OUTDIR, -o OUTDIR
                        output directory for generated SPK
  --overwrite           force overwriting an existing kernel if it exists
  -v, --verbose         set DEBUG level logging output
```

Sub-Command `make-kernel jpss-ck`

```
usage: libera-utils make-kernel jpss-ck [-h] --outdir OUTDIR [--overwrite] [-v] packet_
↪data_filepaths [packet_data_filepaths ...]

positional arguments:
  packet_data_filepaths
                        paths to L0 packet files
```

(continues on next page)

(continued from previous page)

```
options:
  -h, --help            show this help message and exit
  --outdir OUTDIR, -o OUTDIR
                        output directory for generated CK
  --overwrite           force overwriting an existing kernel if it exists
  -v, --verbose         set DEBUG level logging output
```

Sub-Command `make-kernel azel-ck`

```
usage: libera-utils make-kernel azel-ck [-h] [--azimuth] [--elevation] --outdir OUTDIR [-
↪-overwrite] [--csv] [-v] packet_data_filepaths [packet_data_filepaths ...]

positional arguments:
  packet_data_filepaths
                        paths to L0 packet files

options:
  -h, --help            show this help message and exit
  --azimuth             generate ck for Azimuth
  --elevation           generate ck for Elevation
  --outdir OUTDIR, -o OUTDIR
                        output directory for generated CK
  --overwrite           force overwriting an existing kernel if it exists
  --csv                 the provided Az and El packet_data_filepaths are ASCII csv files␣
↪instead of binary CCSDS
  -v, --verbose         set DEBUG level logging output (otherwise set by LIBSDP_STREAM_
↪LOG_LEVEL)
```

Sub-Command `step-function-trigger`

```
usage: libera-utils step-function-trigger [-h] [-w] [-v] algorithm_name applicable_day

positional arguments:
  algorithm_name        Algorithm name you want to run
  applicable_day        Day of data you want to rerun. Format of date: YYYY-MM-DD

options:
  -h, --help            show this help message and exit
  -w, --wait_for_finish
                        Block command line until step function completes (may be a long␣
↪time)
  -v, --verbose         Prints out the result of the step_function_trigger run
```

1.2 File Handling

See the smart_open API documentation here Also see *Working with NetCDF4 Files*

The libera-utils smart_open function has the capability to read and write files to/from a local directory or S3 bucket transparently. It supports a context manager pattern and the usual modes for reading/writing/binary provided by most Python filelike objects:

```
from libera_utils.io.smart_open import smart_open
from libera_utils.io.filenaming import LiberaDataProductFilename

f = LiberaDataProductFilename("s3://some-bucket/LIBERA_L1B_RAD-4CH_V1-2-3_
↪20250102T120000_20250103T120000_R25005112233.nc")
with smart_open(f.path, "r") as filehandler:
    # Work with file contents
    pass
```

1.3 File Naming

The Libera Utils `Filename` classes allow reliable file naming, checking, and path management to support conformity with the Libera filenaming conventions. Each type of filename contains regex that validates every definition or update of the internally tracked filename string. These classes transparently support both S3 paths and local filepaths, including dynamic switching between the two, to simplify the transition between local development environments and AWS.

Full specifics including all available file naming classes are available *in the filenaming API documentation here*

Below is an example test using a `LiberaDataProductFilename` instance to manage a filename string, including switching between S3 and local paths to show the flexibility of the classes.

```
from pathlib import Path
from cloudpathlib import S3Path
from libera_utils.io import filenaming

p = filenaming.LiberaDataProductFilename(
    'LIBERA_L2_CF-RAD_V1-2-3_20270102T112233_20270102T122233_R27002112233.nc')
# Add an S3 prefix
p.path = S3Path('s3://bucket') / p.path
assert isinstance(p.path, S3Path)
# Change prefix to local
p.path = Path('/tmp/path') / p.path.name
assert isinstance(p.path, Path)
# Remove basepath altogether
p.path = p.path.name
assert isinstance(p.path, Path)
# Check that providing a bad value for a basepath doesn't pollute the instance's valid␣
↪path
try:
    p.path = '/bad/prefix' + p.path.name  # The missing / will make this fail regex␣
↪validation
    raise Exception('The previous line should have raised a ValueError')
except ValueError as e:
    assert "failed validation against regex pattern" in str(e)
assert p.path.name == 'LIBERA_L2_CF-RAD_V1-2-3_20270102T112233_20270102T122233_
↪R27002112233.nc'
```

1.4 Creating Libera Data Product Files (NetCDF-4)

TODO[LIBSDC-633]: Include discussion here (or in the YAML format description below) of Dimensions, static global attributes, and dynamic attributes.

The Libera SDC provides utilities for creating NetCDF4 data product files that conform to YAML data product definitions and a few global expectations. These definitions ensure a degree of consistency across all Libera data products, including proper metadata, coordinate systems, and data types. Libera Utils includes a few different ways to work with data product definitions for validation.

1.4.1 Basic Usage

The simplest way to create a valid Libera data product is using the `write_libera_data_product` function:

```
import numpy as np
from pathlib import Path
from libera_utils.io.netcdf import write_libera_data_product

# Define the path to the data product definition YAML file
product_definition_path = Path("path/to/product_definition.yml")

# Create numpy data arrays for each variable and coordinate required
n_times = 100
time_data = np.arange(n_times).astype("datetime64[ns]") + np.datetime64("2024-01-01")
lat_data = np.linspace(-90, 90, n_times)
lon_data = np.linspace(-180, 180, n_times)
fil_rad_data = np.random.rand(n_times) * 1000  # Filtered radiance values
q_flag_data = np.zeros(n_times, dtype=np.int32)  # Quality flags

# Combine all data into a dictionary
data = {
    "time": time_data,
    "lat": lat_data,
    "lon": lon_data,
    "fil_rad": fil_rad_data,
    "q_flag": q_flag_data,
}

# Call write_libera_data_product with correct arguments
output_dir = Path("/path/to/output/directory")
filename = write_libera_data_product(
    data_product_definition=product_definition_path,
    data=data,
    output_path=output_dir,
    time_variable="time",  # Specify which variable contains time data
    strict=True,  # Enforce strict conformance, raising exceptions on errors
)

print(f"Created data product: {filename.path}")
# Output: Created data product: /path/to/output/directory/LIBERA_L1B_RAD-4CH_V0-0-1_
↪20240101T000000_20240101T002739_R25280215327.nc
```

The function automatically:

- Validates data against the product definition
- Generates a standardized filename based on the time range
- Applies correct encoding and compression settings
- Ensures CF-convention compliance

1.4.2 Advanced Usage

There are a few other APIs that may be useful for debugging data product creation during development. Once a data product definition is working, the proper approach is to call `write_libera_data_product` with `strict=True`.

Creation, Enforcement, and Validation for Existing Dataset

You can use the `DataProductDefinition` class to create, enforce (modify), and validate xarray Dataset objects.

Validation of Conformance

The `DataProductDefinition.check_dataset_conformance()` method takes a dataset as an argument and checks that it conforms to the data product definition. The typical usage of this method is in strict mode, which raises an exception for any validation error. For debugging, you can run with `strict=False`, in which case it returns a (possibly empty) list of informational error messages.

```
# Call DataProductDefinition.check_dataset_conformance() with strict=False
# so we get a list of error messages instead of an exception raised.
# This checks the conformance of an existing dataset against a data product definition␣
↪without modifying it
# Note that the dataset need not be generated by create_conforming_dataset. This method␣
↪accepts _any_ xarray Dataset object.
validation_errors = dpd.check_dataset_conformance(dataset, strict=False)

# Demonstrate how to use the return values
if validation_errors:
    print(f"\nValidation found {len(validation_errors)} issues")
    print("Attempting to fix issues with enforce_dataset_conformance...")
```

Enforcement of Product Definition

The `DataProductDefinition.enforce_dataset_conformance()` method takes a Dataset as an argument and attempts to enforce all data product definition requirements by modifying any necessary attributes, data types, and encodings. This doesn't necessarily work, so we return a list of error messages for any problems that couldn't be fixed.

```
# Call dpd.enforce_dataset_conformance() to fix problems with a dataset
# This modifies the dataset in-place to fix what it can
# Obviously some problems are not fixable such as missing variables.
fixed_dataset, remaining_errors = dpd.enforce_dataset_conformance(dataset)

# Demonstrate how to use the return values
if not remaining_errors:
    print("Successfully fixed all conformance issues!")
    # Save the fixed dataset
    fixed_dataset.to_netcdf("output_fixed.nc")
else:
    print("Some issues could not be automatically fixed")
    print(f"Remaining issues: {len(remaining_errors)}")
    for error in remaining_errors:
        print(f"  - {error}")
```

Direct Creation of Conforming Dataset

```
import numpy as np
import xarray as xr
from pathlib import Path
from libera_utils.io.product_definition import DataProductDefinition

# Create DataProductDefinition object from a YAML file
definition_path = Path("path/to/product_definition.yml")
dpd = DataProductDefinition.from_yaml(definition_path)

# Prepare your data as numpy arrays
n_times = 50
data = {
    "time": np.arange(n_times).astype("datetime64[ns]") + np.datetime64("2024-03-01"),
    "lat": np.linspace(-90, 90, n_times),
    "lon": np.linspace(-180, 180, n_times),
    "fil_rad": np.random.rand(n_times) * 500,
    "q_flag": np.random.randint(0, 100, n_times, dtype=np.int32),
}

# Create a Dataset with dpd.create_conforming_dataset() with strict=False
# This allows creation even if some requirements aren't met.
# This method coerces the resulting dataset to conform as closely as possible
# to the data product definition by adding attributes, encodings,
# and coercing data types if possible.
dataset, errors = dpd.create_conforming_dataset(
    data=data,
    strict=False  # Don't raise exceptions, just report issues
)

# Demonstrate how to use the return values
if not errors:
    print("Dataset is valid and conforms to the product definition")
else:
    print(f"Dataset has {len(errors)} conformance issues:")
    for error in errors[:5]:  # Show first 5 errors
        print(f"  - {error}")
```

Dynamic Global and Variable Attributes

Note: L2 Developers should not need this.

Avoid this unless absolutely necessary!

When creating datasets, you can specify attributes whose values are not defined in the YAML (i.e. they are null). We call these "dynamic attributes" and they are required but must be passed directly during dataset creation:

```
# Set dynamic global attribute values (e.g., algorithm-specific metadata values for␣
↪required attribute keys)
user_global_attrs = {
    "date_created": "2024-03-15T11:22:33",
    "algorithm_version": "2.1.0",
    "processing_level": "L1B",
```

(continues on next page)

(continued from previous page)

```
}

# Set dynamic variable attribute values
# NOTE: these should be avoided unless absolutely necessary
user_var_attrs = {
    "fil_rad": {
        "calibration_date": "2024-01-01",
        "sensor_id": "LIBERA-001",
    },
    "q_flag": {
        "flag_meanings": "good questionable bad",
        "flag_values": [0, 1, 2],
    },
}

# Create dataset with custom attributes
dataset, errors = dpd.create_conforming_dataset(
    data=data,
    user_global_attributes=user_global_attrs,
    user_variable_attributes=user_var_attrs,
    strict=True,
)
```

1.4.3 Product Definition YAML Structure

TODO[LIBSDC-623]: Include discussion of dimension names being in a global namespace.

Product definition files specify the exact structure and metadata for data products:

```
# Product level metadata attributes
attributes:
  ProductID: RAD-4CH
  version: 0.0.1

# Coordinate variables (dimensions)
coordinates:
  time:
    dtype: datetime64[ns]
    dimensions: ["time"]
    attributes:
      long_name: "Time of sample collection"
    encoding:
      units: nanoseconds since 1958-01-01
      calendar: standard
      dtype: int64

# Data variables
variables:
  fil_rad:
    dtype: float64
    dimensions: ["time"]
    attributes:
```

(continues on next page)

(continued from previous page)

```
    long_name: Filtered Radiance
    units: W/(m^2*sr*nm)
    valid_range: [0, 1000]
```

The DataProductDefinition system ensures all Libera data products maintain consistent structure, metadata, and quality standards across the mission.

1.5 Making and Using Manifest Files

All science algorithms that run on the Libera Science Data Center system need capabilities for dealing with Manifest Files. Specifics on the usage of manifest files can be found in the *Manifest API documentation here*

The `Manifest` class is designed to handle reading, writing, and interacting with manifest files during processing. It performs such tasks as validating manifest file structure and naming conventions as well as storing the manifest contents as easily accessible python objects and providing helper methods for common tasks related to manifest file handling.

```
from libera_utils.io.manifest import Manifest

# Manifest filenames are passed into your Docker image CLI as its only argument
input_manifest = Manifest.from_file("s3://some-dropbox/LIBERA_INPUT_MANIFEST_
↪01H2GK8J6XM93VKQP4CQFM1TAN.json")
# Read from manifest file to do processing

# Create an output manifest named according to the input manifest (timestamp matches for␣
↪traceability)
output_manifest = Manifest.output_manifest_from_input_manifest(input_manifest)

# Add files. This will raise a credentials error because it tries to checksum the file␣
↪but can't access S3
# without credentials provided (your Docker images will have proper credentials␣
↪attached).
output_manifest.add_files(
    "s3://some-dropbox/LIBERA_L2_CF-RAD_V1-2-3_20270102T112233_20270102T122233_
↪R27002112233.nc"
)

# Automatically generates a proper output manifest filename and writes it to the path␣
↪specified,
# usually this path is retrieved from the environment, like `os.environ["PROCESSING_PATH
↪"]`.
output_manifest.write("s3://some-dropbox/")
```

1.6 Logging

High quality logging is an important part of operational processing and the Libera SDC Team has made logging setup as painless as possible, while also offering a high degree of configuration for processing algorithms. See below for a general discussion of logging principles followed by some example use cases.

See the `libera_utils.logutil` API documentation here

1.6.1 Logging vs. `print`

Printing is a valid way to log from your code. However, it is limited in a few major ways:

1. You get no logging information from library code you have pulled in as dependencies.
2. There is no easy way to automate adding context to print statements such as current function, line, module, etc.
3. Formatting is only a convention with print calls, which makes log analysis and monitoring difficult.
4. There is no easy way to control the verbosity of your print statements.
5. You can only send messages to the console (stdout/stderr).

When using the Libera Utils logging module, you get:

1. Fine-grained information from all the libraries you are using.
2. Configurable standard context added to logs such as time, severity level, module, line number, function name, etc.
3. Consistent formatting to make logs easily searchable.
4. Ability to easily turn logging on/off from one place in the code.
5. Send log messages to multiple configurable destinations (console, file, etc).

See examples of these use cases in the code examples throughout this page.

1.6.2 Logging Levels in Python

- DEBUG - Detailed information, typically of interest only when diagnosing problems.
- INFO - Confirmation that things are working as expected.
- WARNING - An indication that something unexpected happened, or indicative of some problem in the near future (e.g. ‘disk space low’). The software is still working as expected.
- ERROR - Due to a more serious problem, the software has not been able to perform some function.
- CRITICAL - A serious error, indicating that the program itself may be unable to continue running.

1.6.3 Setting Up Logging in Applications

The `libera_utils.logutil` module provides utilities for configuring logging easily for file-based, stream (stdout/stderr), and custom AWS CloudWatch logging.

Simplest Logging Setup

This example allows all messages through to the console at the specified level. This does not filter out DEBUG logs from verbose libraries.

```
"""Simplest logging setup"""
import logging
from datetime import datetime, timezone
import boto3
from botocore.exceptions import NoCredentialsError
from libera_utils.logutil import configure_task_logging

logger = logging.getLogger(__name__)

if __name__ == "__main__":
```

(continues on next page)

(continued from previous page)

```
    task_id = f'processing-task-{datetime.now(tz=timezone.utc).strftime("%Y-%m-%dT%H:%M:
↪%S")}'
    configure_task_logging(task_id, console_log_level="DEBUG")

    logger.debug("test debug message")

    try:
        # The following will demonstrate why we might want to filter out debug messages
        buckets = boto3.client('s3').list_buckets()
        logger.info(buckets)
    except NoCredentialsError:
        logger.error("No credentials found")
```

produces

```
2024-04-23 07:54:37,523 INFO     [libera_utils.logutil:logutil.py:257 in configure_task_
↪logging()]: Console logging configured at level DEBUG.
2024-04-23 07:54:37,523 DEBUG    [__main__:scratch_11.py:15 in <module>()]: test debug␣
↪message
2024-04-23 07:54:37,523 DEBUG    [botocore.hooks:hooks.py:482 in _alias_event_name()]:␣
↪Changing event name from creating-client-class.iot-data to creating-client-class.iot-
↪data-plane
2024-04-23 07:54:37,524 DEBUG    [botocore.hooks:hooks.py:482 in _alias_event_name()]:␣
↪Changing event name from before-call.apigateway to before-call.api-gateway
2024-04-23 07:54:37,524 DEBUG    [botocore.hooks:hooks.py:482 in _alias_event_name()]:␣
↪Changing event name from request-created.machinelearning.Predict to request-created.
↪machine-learning.Predict
2024-04-23 07:54:37,524 DEBUG    [botocore.hooks:hooks.py:482 in _alias_event_name()]:␣
↪Changing event name from before-parameter-build.autoscaling.CreateLaunchConfiguration␣
↪to before-parameter-build.auto-scaling.CreateLaunchConfiguration
2024-04-23 07:54:37,525 DEBUG    [botocore.hooks:hooks.py:482 in _alias_event_name()]:␣
↪Changing event name from before-parameter-build.route53 to before-parameter-build.
↪route-53
2024-04-23 07:54:37,525 DEBUG    [botocore.hooks:hooks.py:482 in _alias_event_name()]:␣
↪Changing event name from request-created.cloudsearchdomain.Search to request-created.
↪cloudsearch-domain.Search
2024-04-23 07:54:37,525 DEBUG    [botocore.hooks:hooks.py:482 in _alias_event_name()]:␣
↪Changing event name from docs.*.autoscaling.CreateLaunchConfiguration.complete-section␣
↪to docs.*.auto-scaling.CreateLaunchConfiguration.complete-section
2024-04-23 07:54:37,526 DEBUG    [botocore.hooks:hooks.py:482 in _alias_event_name()]:␣
↪Changing event name from before-parameter-build.logs.CreateExportTask to before-
↪parameter-build.cloudwatch-logs.CreateExportTask
2024-04-23 07:54:37,526 DEBUG    [botocore.hooks:hooks.py:482 in _alias_event_name()]:␣
↪Changing event name from docs.*.logs.CreateExportTask.complete-section to docs.*.
↪cloudwatch-logs.CreateExportTask.complete-section
2024-04-23 07:54:37,526 DEBUG    [botocore.hooks:hooks.py:482 in _alias_event_name()]:␣
↪Changing event name from before-parameter-build.cloudsearchdomain.Search to before-
↪parameter-build.cloudsearch-domain.Search
2024-04-23 07:54:37,526 DEBUG    [botocore.hooks:hooks.py:482 in _alias_event_name()]:␣
↪Changing event name from docs.*.cloudsearchdomain.Search.complete-section to docs.*.
↪cloudsearch-domain.Search.complete-section
...and much much more
```

Notice the huge volume of DEBUG messages originating from loggers in the `botocore` *package.*

Filtered Logging Setup

Now we want to alter the code above to take advantage of the ability to reduce the amount of DEBUG spam from libraries that we are not interested in. Note the use of `__main__` in the `limit_debug_loggers` tuple. This allows debug messages that originate at the level of a runnable python script that is wrapped in the standard `if __name__=="__main__"` guard.

```
"""Simplest logging setup"""
import logging
from datetime import datetime, timezone
import boto3
from botocore.exceptions import NoCredentialsError
from libera_utils.logutil import configure_task_logging

logger = logging.getLogger(__name__)

if __name__ == "__main__":
    task_id = f'processing-task-{datetime.now(tz=timezone.utc).strftime("%Y-%m-%dT%H:%M:
↪%S")}'
    configure_task_logging(task_id, limit_debug_loggers=("__main__", "libera_utils"),␣
↪console_log_level="DEBUG")

    logger.debug("test debug message")

    try:
        # The following will demonstrate why we might want to filter out debug messages
        buckets = boto3.client('s3').list_buckets()
        logger.info(buckets)
    except NoCredentialsError:
        logger.error("No credentials found")
```

produces

```
2024-04-23 07:52:56,009 INFO     [libera_utils.logutil:logutil.py:257 in configure_task_
↪logging()]: Console logging configured at level DEBUG.
2024-04-23 07:52:56,009 DEBUG    [__main__:scratch_11.py:15 in <module>()]: test debug␣
↪message
2024-04-23 07:52:56,186 ERROR    [__main__:scratch_11.py:22 in <module>()]: No␣
↪credentials found
```

Contrived Runnable Example

This illustrates how the `limit_debug_loggers` kwarg works for preventing other libraries from logging at DEBUG level while still allowing DEBUG messages from libraries you care about.

```
"""Logging setup contrived example"""
from datetime import datetime, timezone
import logging
from libera_utils.logutil import configure_task_logging

task_id = f'processing-task-{datetime.now(tz=timezone.utc).strftime("%Y-%m-%dT%H:%M:%S")}
↪'
```

(continues on next page)

(continued from previous page)

```
configure_task_logging(task_id, limit_debug_loggers=("my_package",), console_log_
↪level=logging.DEBUG)
# my_log is a logger from inside your library (name prefixed with your library name).
my_log = logging.getLogger('my_package.my_application')
# The following is an example. External libraries will create their own loggers␣
↪internally.
external_library_log = logging.getLogger('some_spammy_library')

my_log.debug('subtle but important message gets passed through')
external_library_log.debug('this external library debug spam gets filtered out')
external_library_log.info('and external library info messages still get through')
```

Configuring Stream Logging

Stream logging needs only a level and it defaults to `INFO`. Stream logging cannot be disabled but is easily ignored.

Note: You can change the default formatter for stream logging from plaintext to json by passing `console_log_json=True` to `configure_task_logging`. This is convenient for logging in AWS services that push their stdout logs to CloudWatch.

```
"""Example of setting up console logging (JSON formatted)"""
import logging
from libera_utils.logutil import configure_task_logging

configure_task_logging("test-task-id-1", console_log_json=True, console_log_
↪level=logging.DEBUG)
```

Configuring Filesystem Logging

Filesystem logging needs only a log directory (it always logs at DEBUG level). If you don't pass `log_dir`, no file-based logging will occur. The log directory must exist and will not be dynamically created.

Note: Logging to a directory is really only useful for local testing. Any logs written to a directory in a docker container will evaporate upon completion of the docker container process.

```
"""Example of setting up file-based logging"""
from pathlib import Path
from libera_utils.logutil import configure_task_logging

configure_task_logging("test-task-id-1", log_dir=Path("/tmp"))
```

1.6.4 Logging Exceptions

The Python logging module provides logging calls associated with each level. In addition, it provides a logging call for logging exceptions that includes the current stack trace for debugging.

```
"""Example logging calls"""
import logging

logger = logging.getLogger(__name__)

if __name__ == "__main__":
    logging.basicConfig(level=logging.DEBUG)
```

(continues on next page)

(continued from previous page)

```
    logger.debug("test debug")
    logger.info("test info")
    logger.warning("test warning")
    logger.error("test error")
    logger.critical("test critical")

    try:
        raise ValueError("example exception to log")
    except ValueError:
        logger.exception("encountered exception")  # This logs the message followed by␣
↪the exception traceback
        # raise  # Optional re-raise of exception after logging it
```

1.6.5 Concept: Module Level Logging

(AKA library logging)

Module level logging is the practice of defining a logger at the top of each module (.py file) and using that logger object for the entire module. Modules in libera_utils should all have module level loggers when appropriate. e.g.

```
"""Example of module level logger instantiation"""
import logging

logger = logging.getLogger(__name__)

# Module code
```

One advantage of module level logging is that it provides a logger, named for the module from which it is logging but doesn't configure the logger at the module level. Since much of this code is intended to be reused by others, we avoid configuring loggers in reusable code. If a logger is needed in a context, it should be configured at the "application level". That is, the top level application (e.g. CLI tool) that is running a process should take care of logging configuration and assume that each module has generic module level loggers configured for the internal code to use.

Another advantage of module level logging is that our loggers come out with automatically structured names like `libera_utils.db.database` and they all start with `libera_utils`. This allows us to treat those loggers differently than those named, for example, `some_spammy_library.emit_spam`. In our logging setup, we allow users to turn off debug messages for all loggers that aren't named with specific prefixes. This allows us to pass up debug messages from our code but ignore debug messages from dependency code (AWS boto APIs in particular spam a LOT of debug messages).

1.6.6 Fully Customized Logging

If you want complete control over your logging configuration, you can use our provided `configure_static_logging` function, which reads a YAML configuration file that represents a Python logging configuration. This is a completely static configuration and should be supplied as part of your processing algorithm application code.

```
"""Example of configuring logging with static config file"""
import logging
from pathlib import Path
from libera_utils.logutil import configure_static_logging

config = Path("/path/to/config_file.yml")
configure_static_logging(config)
```

(continues on next page)

(continued from previous page)

```
logger = logging.getLogger()
logger.info("handling depends on your supplied configuration")
```

An example of a logging config file:

```
# Example parameterized logging configuration
version: 1
disable_existing_loggers: False
formatters:
  json:
    format: '{"time": "%(asctime)s",
      "level": "%(levelname)s",
      "module": "%(filename)s",
      "function": "%(funcName)s",
      "line": %(lineno)d,
      "message": "%(message)s"}'
  plaintext:
    format: "%(asctime)s %(levelname)-9.9s [%(filename)s:%(lineno)s in %(funcName)s()]:
↪%(message)s"
handlers:
  console:
    class: logging.StreamHandler
    formatter: plaintext
    level: INFO
    stream: ext://sys.stdout
  logfile:
    class: logging.handlers.RotatingFileHandler
    formatter: plaintext
    level: DEBUG
    filename: /tmp/libera_utils_test_log.log
    maxBytes: 1000000
    backupCount: 3
root:
  level: INFO
  propagate: True
  handlers: [console, logfile]
loggers:
  libera_utils:
    qualname: libera_utils
    level: DEBUG
    handlers: []
```

CHAPTER

TWO

DEVELOPER DOCS

This section is for developers working on the Libera Utils package but also includes reference documentation for the tools used for development and how to use them (e.g. git and Docker).

2.1 Setting Up a Development Environment

2.1.1 Managing Multiple Base Python Versions

In order to develop with multiple different versions of Python and create virtual environments associated with different versions of Python, you will need multiple base Python interpreters. There are several ways to manage this including Conda, PyEnv, and building Python from source. We recommended using Conda and outline the steps below for using Conda to manage multiple base Python installations.

1. Install miniconda according to the official documentation. If you already have miniconda or anaconda installed, you can skip this step.
2. Optionally run `conda config --set auto_activate_base false` to add a configuration to your `.condarc` file to disable auto-activation of the `base` conda environment on shell startup.
3. Create a conda environment with your preferred version of python: `conda create -n conda-python3.11 python=3.11`
    - Note: Name this environment with a convention that makes sense to you for a base interpreter. *Do not delete this conda environment!* Deleting it will break all subsequent virtual environments based on it.
4. The Python interpreter provided by your new conda environment is a full base interpreter and you can use it to create virtual environments. You can find the full path to the base interpreter by running something similar to the following (run `conda env list` to see why this works):

```
PATH_TO_PYTHON=$(conda env list | grep "conda-python3.11" | awk '{print $2}')/bin/
↪python
$PATH_TO_PYTHON -m venv path/to/new/venv
```

2.1.2 Installing Poetry

Poetry is a command line tool that helps manage a python development environment, including package management, virtual environment management, and package building.

Poetry official installation instructions can be found here: https://python-poetry.org/docs/#installation. We recommend using the installation script provided by poetry not using pipx.

To ensure that Poetry is always available and in the `PATH` it is recommended to install Poetry with your pre-installed system python interpreter rather than as a package in a conda environment or in a virtual environment. The specific version of python with which you install Poetry is inconsequential (as long as it is currently supported by Poetry). If your system python is not supported by Poetry, you can install Poetry in your conda base environment. Just remember

that Poetry will only be available when that environment is activated. Things can get a bit confusing when you have a conda environment active and a derived virtual environment activated on top of it.

Once poetry is installed, check that it works by running `poetry --version`. You should get something like

```
Poetry version 2.1.0
```

Installing Poetry with System Python

Ensure that all your virtual environments and conda environments are deactivated and that `which python3` refers to your system python interpreter (usually `/usr/bin/python3`).

```
curl -sSL https://install.python-poetry.org | python3 -
```

2.1.3 Configuring Poetry

We recommend creating your own virtual environments in locations of your choosing (common to create them in the project directory, as `venv` or `.venv`).

2.1.4 Setting Up Development Virtual Environment(s)

Poetry will dynamically detect the presence of an activated virtual environment and use that if present. If none is present, Poetry will automatically create one for you in your user cache location (e.g. on mac in `~/Library/Caches/pypoetry/virtualenvs`). This can be confusing so we recommend creating your venv and letting poetry use it when activated.

1. Deactivate all Conda environments and virtual environments
2. Activate the conda environment you wish to use for your base python interpreter (e.g. `conda activate conda-python3.11`)
3. Create a virtual environment anywhere you wish. `python -m venv path/to/venv`.
4. Activate newly created venv: `source path/to/venv/bin/activate`
5. [Recommended]: Configure your IDE to recognize the correct poetry-managed virtual environment for the version you wish to develop with.
6. Run `poetry env info` and verify that Poetry is recognizing your virtual environment properly:

```
Virtualenv
Python:         3.9.9
Implementation: CPython
Path:           /Users/myuser/path/to/libera_utils/venv
Valid:          True
```

Changing Python Versions

It is common to recreate your virtual environment on a regular basis in order to use different python versions. You can do this by making sure that `python` points to the base interpreter you wish to use (e.g. 3.12) and going through the steps above to create a new venv (you can name it differently) and activating it for poetry to use.

Installing Dependencies

1. Run `poetry lock && poetry install` in the same directory as the `pyproject.toml` file. You should see poetry solving the dependency tree and then installing dependencies. This also installs dev group dependencies, as specified in `pyproject.toml`. Lastly you should see it installing the local package.

2. To install optional "extra" dependencies (aka optional dependencies), run `poetry install -E extra_name1 -E extra_name2`. These extra dependencies are specified in `pyproject.toml` under `[tool.poetry.extras]`. Note that any subsequent `poetry install` command without `--extras` will implicitly uninstall any previously installed extras.
3. To install dependency "groups", which may or may not be optional, use the `--with` and `--without` flags for Poetry. e.g. `poetry install --with docgen` will install the dependencies for the optional group "docgen".
4. Verify that the `libera_utils` package was installed correctly by running `libera-utils --version`. This runs the `libera-utils` command line utility that is included in the package. This can also be run with `poetry run libera-utils --version`.
5. Set up `pre-commit` by running `pre-commit install`. This installs the standard git hooks that we use to prevent mistakes before they are committed. Configuration for pre-commit can be found in `.pre-commit-config.yaml`.
6. Next, *go run the tests*.

2.2 Configuration

Configurations are stored in a JSON file `config.json` but we allow overriding those values with environment variables. If, at any point in the code, the config is queried (`config.get(key)`) for a key that isn't present in the top level of the JSON file, a warning is issued to add a default value to the JSON file. This helps ensure that we can track all possible configuration values in the JSON config rather than having to bookkeep them elsewhere.

2.3 Testing

Testing is run with `pytest`. To run all tests, make sure the dev dependencies are installed and run:

```
# Run all unit and integration tests
pytest tests
```

Pytest configuration is stored in `pyproject.toml` under `[tools.pytest.ini_options]`.

Tests are stored in the `tests` directory and are divided into `integration` and `unit` tests. Unit tests check for expected behavior of small pieces of the package (e.g. a single function). Integration tests check for "larger" behavior across wider swaths of the package, testing that components function together cohesively. Integration test modules contain the pytest mark `pytestmark = pytest.mark.integration`. This allows you to selectively exclude running integration tests with

```
pytest -m "not integration" tests
```

Unit tests heavily utilize pytest parameterization in order to run a single test against many sets of inputs and outputs without duplicating code (see [parameterize] (https://docs.pytest.org/en/stable/example/parametrize.html#set-marks-or-test-id-for-individual-parametrized-test)).

We also heavily utilize custom pytest [fixtures] (https://docs.pytest.org/en/stable/how-to/fixtures.html#requesting-fixtures), defined in `tests/plugins`. These plugins are made available to pytest in `conftest.py`, the main test configuration file for pytest.

In order to better ensure test independence, we use `pytest-randomly` to randomize the order of tests. The random seed used to set the order is printed at the beginning of a test run. You can re-run the tests with a specific random seed by running `pytest --randomly-seed=<seed>`.

2.3.1 Generating Coverage and Test Reports

With coverage for generating reports on code coverage:

```
# Create coverage data (stored in .coverage)
pytest --cov=libera_utils --junit-xml=junit.xml
# Generate interactive HTML coverage report
pytest --cov-report=html:coverage_report --cov=libera_utils
# Generate Corbertura-compatible XML report
pytest --cov-report=xml:coverage.xml --cov=libera_utils
```

2.3.2 Test Profiling

Time Profiling

We use the `pytest-profiling` plugin for time profiling of tests.

To create profiling output for a test:

```
pytest --profile tests/test_module.py::test_specific_test[PARAM_ID]
```

This generates a `prof` directory. To visualize the results, we have a tool called `snakeviz` included in dev dependencies.

To visualize profiling results with `snakeviz`:

```
snakeviz prof/combined.prof
```

It will start a web server with a navigable GUI of timing profiles for each part of the stack.

Memory Profiling

We use the `memory-profiler` package for profiling memory usage in tests. Documentation here: https://github.com/pythonprofilers/memory_profiler

2.3.3 Testing in Docker

To run the unit tests in docker, run

```
docker-compose up [--build] --exit-code-from=tests tests --attach=tests
```

This runs the `tests` container service defined in the `docker-compose.yml` file. The `--build` option forces docker to rebuild the testing container image before running (e.g. if things have changed).

Copying Test Report Artifacts from Docker

When we run tests in Docker on Jenkins, we often want to copy and save Corbertura and JUnit test reports. Jenkins has facility for doing this easily with

```
always {
    junit '**/*junit.xml'
    cobertura coberturaReportFile: '**/*coverage.xml'
}
```

The challenge when running in Docker is to make these test artifacts available to Jenkins. By default these files exist only inside the Docker container so we must copy them out. Do this with

```
docker-compose --exit-code-from tests up tests
docker-compose cp tests:/path/to/report.xml .
```

2.3.4 Static Analysis

NASA requirements document NPR7150.2C requires that we perform static analysis of our codebase to check for common vulnerabilities and statically detectable code weaknesses and vulnerabilities (CWEs and CVEs).

We use the Ruff tool to perform a comprehensive static analysis of our code. Ruff includes configurations for pycodestyle, flake8, Bandit, and more. It is configured to run automatically as a pre-commit hook via the pre-commit tool.

To manually run all ruff checks, run

```
ruff check
```

Configuration for ruff is declared in `pyproject.toml`.

2.3.5 Pre-commit Hooks

To ensure code quality with minimal effort on the part of developers, we use pre-commit to run automatic linting before commits are allowed. Configuration for pre-commit is in `.pre-commit-config.yaml`.

To install pre-commit, run:

```
pre-commit install
```

To run all hooks on all files manually, run:

```
pre-commit run --all-files
```

2.4 Testing Another Package Against a Working Version of Libera Utils

When developing Libera Utils, it is often useful to test that package against a specific commit hash of Libera Utils to make sure it works as expected.

Update the `pyproject.toml` file of the dependent project with a reference to the specific git hash of Libera Utils.

```
[tool.poetry.dependencies]
libera_utils = { git = "ssh_or_http_url_to_libera_utils_repo", rev = "abc123def456..." }
```

If your pyproject.toml is using the PEP 621 format:

```
[project]
dependencies = [
    "libera_utils @ git+ssh_or_http_url_to_libera_utils_repo@abc123def456..."
]
```

Then uninstall the old version (if you don't, Poetry sometimes assumes it's already installed if the version isn't different in the specific commit you want to reference) and reinstall dependencies:

```
pip uninstall libera_utils
poetry lock && poetry sync
```

2.5 Build and Release

Our release pipeline is an automated process that allows us to build and release the `libera-utils` package in a consistent manner whenever tags are pushed to the repo. When a tag is pushed, the CI/CD pipeline will automatically build the package and upload it to PyPI with the version number specified in the `pyproject.toml` file.

Note: The version number in `pyproject.toml` is the source of truth for the version of the package, not the tag text. The tag text is used to trigger the CI/CD pipeline, but the version number in `pyproject.toml` is what is used when building and publishing the package to PyPI. The convention is to use the tag text as the version number for consistency.

2.5.1 Release Process for Major and (Usually) Minor Releases

Every PR (with a few exceptions) should include a version bump to the package and an update in to `doc/source/changelog.md`. When a PR is merged, the version should already be updated. Once a PR is merged, one of the lead developers is responsible for tagging the commit in `main` and pushing the tag to trigger the CI pipeline that pushes the release to PyPI.

To tag a commit and push the tag, use an "annotated" tag as follows:

```
git tag -a X.Y.Z -m "version release X.Y.Z"  # -a indicates annotated, -m is the␣
↪annotation message
git push origin X.Y.Z
```

2.5.2 Release Process for Pre-Releases and Testing in Dependent Packages

Note: This should rarely be necessary. See *testing.md* for how to test other packages against a specific commit hash of Libera Utils without using a pre-release on PyPI.

1. In your working branch, bump the version of the package to the next pre-release version (e.g. `X.Y.Zrc1`) by editing `pyproject.toml`. Make sure that the version you are specifying doesn't already exist in PyPI.
2. Commit the version change.
3. Tag your working branch with the pre-release version, e.g. `X.Y.Zrc1`, and push the tag upstream. This will trigger the automatic build and publish by Jenkins.

   ```
   git tag -a X.Y.Zrc1 -m "pre-release X.Y.Zrc1"
   git push origin X.Y.Zrc1
   ```

2.6 Documentation Generation for Libera Utils

2.6.1 Creating Documentation with Sphinx

Sphinx documentation is a python library for generating documentation from docstring comments throughout a codebase in combination with other rst or md pages separate from the generated documentation. We use it to generate automatic API documentation for the codebase as well as building developer documentation and user documentation from markdown and restructured text documents (such as this one).

You will need `make` installed on your machine in order to build sphinx docs.

Compatible Comments and Docstrings

Docstrings in this project are expected in the Numpy docstring format.

Writing Custom Documentation Pages

If you wish to add pages that are not part of the generation from the code such as reference pages, use the following steps.

1. Create your files and folders and add them to the project folder `doc/source/`
2. Edit the `doc/source/index.rst` file to add the file you have just created to the toctree (table of contents tree)
3. If you added developer documentation, instead of editing the `index.rst` edit the `doc/source/developer-docs.rst` document or create a new folder and associated rst document containing a toctree that includes your new file.

The folder structure is an organizational convention but the actual page structure comes from the toctree structure in the rst files (note that these documents could also be markdown but rst has better support for TOC listings).

Note: This project is configured to read and interpret both markdown (.md) and reStructured Text (.rst) documents. The following are two basic guides to writing proper comments that will create well formatted outputs.

- reStructured Text
- Markdown

If you have need to translate between rst and markdown, use `pandoc`.

2.6.2 Building HTML Documentation Locally

This should all be done while in the virtual environment that is configured for this project. See the *dev environment setup documentation* for poetry instructions.

1. Run `poetry update --with docgen`
   - This ensures that the poetry install is up-to-date, including the "docgen" dependency group.
2. Run `poetry install --with docgen`
   - This installs all project dependencies, including the `pyproject.toml` docgen group (e.g. Sphinx, etc.)
   - This also ensures that the project itself is installed with its most recent version (as defined in `pyproject.toml`)
3. Navigate to the Sphinx document source folder `cd ./doc`
4. Build the html files in the build folder using the make command `make html`
5. Open the generated `build/index.html` file to go to the documentation homepage.

Oneliner for an IDE configuration:

```
# Must be run from the `doc` working directory
make html && open build/html/index.html
```

2.6.3 Automatic Documentation Publishing with ReadTheDocs

We publish our documentation using ReadTheDocs. Our configuration file for readthedocs is located in `.readthedocs.yaml`. The Libera SDC team has a ReadTheDocs account containing an organization and each team member has a login to view build status and private/hidden doc builds. Only versions set to Public are visible without a login.

ReadTheDocs responds to webhooks sent by both Bitbucket and Jenkins.

Webhook from Bitbucket

The Bitbucket webhook for ReadTheDocs goes directly to ReadTheDocs to trigger a build.

Webhook from Jenkins

ReadTheDocs expects specific payload structure to its webhook API. When Bitbucket sends a payload, it triggers a build for the default branch (`main`) because the structure isn't supported by ReadTheDocs. In order to get specific branch builds, we send a webhook for PR updates to Jenkins, which reformats the payload into the correct structure for ReadTheDocs, which triggers a build for the specific PR branch being modified.

Publishing Official Release Docs

ReadTheDocs automatically builds and publishes new versions for tags, but it requires a trigger to re-fetch any new tags. This means a build must be triggered after the release process tag is applied for a new official docs version to be built.

To do this, log in to ReadTheDocs and click "Add a New Version", then at the bottom, click "Resync Versions". This will trigger ReadTheDocs to fetch the new tag from Bitbucket and build it as an official version.

2.7 Git Usage

2.7.1 Basic Workflow

We use trunk-based development as our basic git workflow. The following is a general order of events:

1. Create a branch from `main` for a feature. Name it based on your feature or ticket. e.g. `feature/LIBSDC-XXX-add-something-cool`
2. Add commits to your feature branch.
3. (Optional) Put up a Draft PR to signify a work in progress while still allowing team members visibility. This also allows automated tests to run based on changes to the PR.
    - Jenkins tests run for draft PRs
    - Only add reviewers if you really expect a thorough review
4. Ensure your branch is rebased and you are satisfied with the state of your code. That is, you don't plan to add any more commits until it's reviewed.
5. Put up a PR to merge into `main`.
6. Wait for review, address all comments, and wait for all reviewers to approve.
7. Merge using the "fast-forward only" strategy in Bitbucket. If properly squashed, should only add 1 commit.
    - In most cases you should squash your branch to a single commit
    - Commit messages merged into `main` should be nicely formatted but we don't have a strict standard.

Suggested Commit Message Formatting

```
Title of Commit

Narrative description of the changes in the commit, limited to 80 character
line length by convention.

You can include bullet points that, for example, summarize specific individual changes:
```

(continues on next page)

(continued from previous page)

```
- MAINT: Maintenance change
- CONFIG: Configuration change
```

2.7.2 Reasons for Rebasing

Rebasing allows us to keep our git history completely linear and avoids creating unnecessary merge commits. Merge commits are dangerous because git doesn't actually know what the code is doing and can "resolve" merge conflicts in ways that result in broken code. No commit in `main` should ever contain broken code if we can help it.

2.7.3 Reasons for Squashing

Squashing reduces the total number of commits in the repository to ~1 per pull request. Since each commit contains a full snapshot of the codebase, this drastically reduces the amount of storage necessary to host our git repo and makes life slightly easier for our beloved Web Team.

2.7.4 Release Branches

Release branches are created in preparation for a tagged release from `main`. They are really just special feature branches that allow us to ensure the repo metadata (version, dependencies, changelog, docs, etc.) are in good shape for a formal release. When we merge a release branch, we delete it and perform the final release from the latest commit in `main` by tagging that commit with the version and pushing this tag which automatically deploys build artifacts to PyPI. See the *build and release docs* for more details on our release process.

2.8 Git LFS (Large File Storage) Usage

We use Git LFS to store large files in a way that doesn't blow up the size of our repo on the git server. Usually each commit contains a snapshot of the repository at that point in time. If you store a 100MB file, your entire repo size will be 100MB * number of commits, which can easily balloon to a big number. Incidentally, this is also why we squash PRs to reduce the number of commits in the main branch over time.

Git LFS Documentation

Tutorial on Using Git LFS

Atlassian Docs on Git LFS

2.8.1 Set Up

Install Git LFS according to the Git LFS official documentation (linked above).

Run the following to initialize Git LFS for your user:

```
git lfs install
```

2.8.2 A Note on Git LFS Authentication and Git GUIs

Git LFS authenticates separately from Git. Historically, Git LFS supported only HTTP auth, which was a huge pain because best practice is to use SSH to authenticate with Git servers (note that GitHub has actually deprecated HTTP authentication). This situation meant that even if you were using SSH for git auth, you still had to provide and store HTTP credentials for Git LFS.

However, as of Git LFS v3.0, SSH authentication is supported!

Git GUIs will require you to configure authentication for both Git and Git LFS. It is recommended to use SSH for both but the configuration processes for GUI programs (and IDEs) are different so we're not addressing it here. GLHF!

2.8.3 Tracking New Files in LFS

LFS keeps track of which files are stored in LFS via the `.gitattributes` file.

To track specific files:

```
git lfs track "<pattern>"  # The double quotes matter to prevent shell expansion
# e.g. track all files in every directory named test_data
git lfs track "**/test_data/*"
# This ^ grabs all the filepaths that match and writes them directly into .gitattributes
```

To track files based on a pattern in `.gitattributes`:

```
# .gitattributes
# Track all netCDF files that live anywhere inside a test_data directory
**/test_data/**/*.nc
```

Documentation on Pattern Syntax

2.8.4 Tracking Existing Files in LFS

If you have already committed a file and you wish to move that file to Git LFS, you can:

Add it specifically using `git lfs track` e.g.

```
git lfs track my_large_file.big
```

This method appears to have some magic sugar behind it that automatically removes and re-adds the file to git tracking.

Alternatively, you can add the appropriate pattern to `.gitattributes` and then remove and re-add the file to git history so Git LFS picks it up.

```
echo "**/*.big" >> .gitattributes  # Add pattern to .gitattributes
git rm --cached my_tracked_file.big  # Remove file from git tracking
git add my_tracked_file.big  # Git LFS should pick it up at this point
```

This method is preferable if you want your files generally tracked by pattern rather than individually.

NOTE: As a bit of a trick, you can combine the above strategies by running `git lfs track` on the pattern you wish to store in Git LFS. Then replace the individual records added to `.gitattributes` with the appropriate generic pattern that matches the specific files to be tracked.

2.8.5 Useful Git LFS Commands

List all files in the current git ref (branch, commit, tag, etc.) currently managed by Git LFS:

```
git lfs ls-files
```

Update the files in your `.git/lfs` directory with the version for your current ref:

```
git lfs fetch
```

Convert local pointer files to full files (from `.git/lfs` directory):

```
git lfs checkout
```

Combine fetch and pull into one step:

```
git lfs pull
```

2.9 Usage of SPICE

Fun fact: SPICE stands for "Spacecraft Planet Instrument C-matrix Events," which were the original primary concerns of those developing the library. The "C" stands for Camera because early instruments using SPICE were cameras.

2.9.1 Static Kernels Generated at Libera SDC

These kernels are part of the package data, are manually edited, and change rarely.

Frame Kernel (FK)

e.g. `libera_fk_v01.tf`

Contains reference frame definitions for JPSS and Libera.

Spacecraft Clock Kernel (SCLK)

e.g. `jpss_sclk_v01.tsc`

Contains specification of the spacecraft clock on JPSS.

Instrument Kernel (IK)

e.g. `libera_cam_ik_v01.ti`, `libera_rad_ik_v01.ti`

Contains geometry specification data of the Libera instruments.

2.9.2 Dynamic Kernels Generated at Libera SDC

These kernels are binary generated kernels. They are ancillary data products and are created as part of pipeline processing.

JPSS Ephemeris Kernel (SPK)

e.g. `libera_jpss_20210408t235850_20210409t015849_vM3m14p159_r25365125959.bsp`

Contains ephemeris data – coordinates in ITRF93 frame – for the JPSS spacecraft body.

JPSS Attitude Kernel (CK)

e.g. `libera_jpss_20210408t235850_20210409t015849_vM3m14p159_r25365125959.bc`

Contains attitude data – quaternion rotations from J2000 – for the JPSS spacecraft body.

Azimuth Rotation Mechanism Attitude Kernel (CK)

e.g. `libera_azrot_20210408t235850_20210409t015849_vM3m14p159_r25365125959.bc`

Contains attitude data for the Libera Azimuth Rotation mechanism.

Elevation Scan Mechanism Attitude Kernel (CK)

e.g. `libera_elscan_20210408t235850_20210409t015849_vM3m14p159_r25365125959.bc`

Contains attitude data for the Libera Elevation Scan mechanism.

2.9.3 Kernels Retrieved from NAIF

These kernels are generated at NAIF. We download them as needed using the `KernelFileCache` class, which is configured with a NAIF index page URL and a regex string that finds the proper download URL on the index page. The downloaded file is put in a cache so the download only occurs after the cached data is of a specified age. If we want to effectively cache some kernels indefinitely, we can put them in an S3 bucket and retrieve them from there instead of from the NAIF website.

Leapseconds Kernel (LSK)

e.g. `naif0012.tls`

Contains leap second data used by time conversion routines.

Development Ephemeris Kernel (SPK)

e.g. `de440s.bsp`

Contains ephemeris data for planetary bodies. The version with "s" appended to the filename covers only more recent time (starts in the 1500s) to reduce filesize.

High Precision Earth Binary Planetary Constants Kernel (PCK)

e.g. `earth_000101_211220_210926.bpc`

Contains high precision orientation data for Earth in the ECEF ITRF93 reference frame.

ITRF93 is a more precise version of the standard IAU_EARTH reference frame provided in the text PCK below.

This kernel is regenerated several times per week so we should retrieve this from NAIF for every processing run.

ITFR93 Reference Frame Kernel for Earth

e.g. `earth_assoc_itrf93.tf`

Used to designate ITRF93 as the default body-fixed frame associated with the Earth.

Standard Text Planetary Constants Kernel (PCK)

e.g. `pck00010.tpc`

Contains orientation data and other planetary constants for planetary bodies.

2.10 Libraries Used for Kernel Handling

We use Curryer for kernel generation and SpiceyPy for generic kernel handling and calls to CSPICE (Curryer also uses SpiceyPy internally).

CHAPTER

THREE

API

This is the API documentation for the entire Libera Utils library.

libera_utils	libera_utils high-level package initialization for common utilities.

3.1 libera_utils

libera_utils high-level package initialization for common utilities.

class libera_utils.**LiberaDataProductDefinition**(** (Keyword-only parameters separator (PEP 3102))*, *coordinates: dict[str,* LiberaVariableDefinition*]*, *variables: dict[str,* LiberaVariableDefinition*]*, *attributes: dict[str, Any]*)

Pydantic model for a Libera data product definition.

Used for validating existing data product Datasets with helper methods for creating valid Datasets and DataArrays.

data_variables

A dictionary of variable names and their corresponding LiberaVariable objects, which contain metadata and data.

Type
dict[str, LiberaVariable]

product_metadata

The metadata associated with the data product, including dynamic metadata and spatio-temporal metadata.

Type
ProductMetadata | None

Attributes

dynamic_attributes
Return product-level attributes that are dynamically defined (null values) in the data product definition

model_extra
Get extra fields set during validation.

model_fields_set
Returns the set of fields that have been explicitly set on this model instance.

static_attributes
Return product-level attributes that are statically defined (have values) in the data product definition

Methods

check_dataset_conformance(dataset[, strict])	Check the conformance of a Dataset object against a DataProductDefinition
copy(*[, include, exclude, update, deep])	Returns a copy of the model.
create_conforming_dataset(data[, ...])	Create a Dataset from numpy arrays that is valid against the data product definition
enforce_dataset_conformance(dataset)	Analyze and update a Dataset to conform to the expectations of the DataProductDefinition
from_yaml(product_definition_filepath)	Create a DataProductDefinition from a Libera data product definition YAML file.
generate_data_product_filename(utc_start, ...)	Generate a standardized Libera data product filename.
model_construct([_fields_set])	Creates a new instance of the *Model* class with validated data.
model_copy(*[, update, deep])	!!! abstract "Usage Documentation"
model_dump(*[, mode, include, exclude, ...])	!!! abstract "Usage Documentation"
model_dump_json(*[, indent, ensure_ascii, ...])	!!! abstract "Usage Documentation"
model_json_schema([by_alias, ref_template, ...])	Generates a JSON schema for a model class.
model_parametrized_name(params)	Compute the class name for parametrizations of generic classes.
model_post_init(context, /)	Override this method to perform additional initialization after *__init__* and *model_construct*.
model_rebuild(*[, force, raise_errors, ...])	Try to rebuild the pydantic-core schema for the model.
model_validate(obj, *[, strict, extra, ...])	Validate a pydantic model instance.
model_validate_json(json_data, *[, strict, ...])	!!! abstract "Usage Documentation"
model_validate_strings(obj, *[, strict, ...])	Validate the given object with string data against the Pydantic model.

construct
dict
from_orm
json
parse_file
parse_obj
parse_raw
schema
schema_json
update_forward_refs
validate

_check_dataset_attrs(*dataset_attrs: dict[str, Any]*) → list[str]

Validate the product level attributes of a Dataset against the product definition

Static attributes must match exactly. Some special attributes have their values checked for validity.

Parameters
dataset_attrs (*dict[str, Any]*) – Dataset attributes to validate

Returns
List of error messages describing problems found. Empty list if no problems.

Return type
list[str]

static _get_static_project_attributes(*file_path=None*)

Loads project-wide consistent product-level attribute metadata from a YAML file.

These global attributes are expected on every Libera data product so we store them in a global config.

Parameters
file_path (*Path*) – The path to the global attribute metadata YAML file.

Returns
Dictionary of key-value pairs for static product attributes.

Return type
dict

classmethod _set_attributes(*raw_attributes: dict[str, Any]*) → dict[str, Any]

Validates product level attributes and adds requirements for globally consistent attributes

Parameters
raw_attributes (*dict[str, Any]*) – The attributes specification in the product definition.

Returns
The validated attributes dictionary, including standard defaults that we always require.

Return type
dict[str, Any]

check_dataset_conformance(*dataset: Dataset*, *strict: bool = True*) → list[str]

Check the conformance of a Dataset object against a DataProductDefinition

This method is responsible only for finding errors, not fixing them.

Parameters

- **dataset** (*Dataset*) – Dataset object to validate against expectations in the product configuration
- **strict** (*bool*) – Default True. Raises an exception for nonconformance.

Returns
List of error messages describing problems found. Empty list if no problems.

Return type
list[str]

create_conforming_dataset(*data: dict[str, ndarray]*, *user_product_attributes: dict[str, Any] | None = None*, *user_variable_attributes: dict[str, dict[str, Any]] | None = None*, *strict: bool = True*) → tuple[Dataset, list[str]]

Create a Dataset from numpy arrays that is valid against the data product definition

Parameters

- **data** (*dict[str, np.ndarray]*) – Dictionary of variable/coordinate data keyed by variable/coordinate name.

- **user_product_attributes** (*dict[str, Any] | None*) – *Algorithm developers should not need to use this kwarg.* Product level attributes for the data product. This allows the user to specify product level attributes that are required but not statically specified in the product definition (e.g. the algorithm version used to generate the product)
- **user_variable_attributes** (*dict[str, dict[str, Any]] | None*) – *Algorithm developers should not need to use this kwarg.* Per-variable attributes for each variable's DataArray. Key is variable name, value is an attributes dict. This allows the user to specify variable level attributes that are required but not statically defined in the product definition.
- **strict** (*bool*) – Default True. Raises an exception for nonconformance.

Returns
Tuple of (Dataset, error_messages) where error_messages contains any validation problems. Empty list if the dataset is fully valid.

Return type
tuple[Dataset, list[str]]

Notes

- We make no distinction between coordinate and data variable input data and assume that we can determine which is which based on coordinate/variable names the product definition.
- This method is not responsible for primary validation or error reporting. We call out to check_dataset_conformance at the end for that.

property dynamic_attributes

Return product-level attributes that are dynamically defined (null values) in the data product definition

These attributes are _required_ but are expected to be defined externally to the data product definition

enforce_dataset_conformance(*dataset: Dataset*) → tuple[Dataset, list[str]]

Analyze and update a Dataset to conform to the expectations of the DataProductDefinition

This method is for modifying an existing xarray Dataset. If you are creating a Dataset from scratch with numpy arrays, consider using *create_conforming_dataset* instead.

Parameters
dataset (*Dataset*) – Possibly non-compliant dataset

Returns
Tuple of (updated Dataset, error_messages) where error_messages contains any problems that could not be fixed. Empty list if all problems were fixed.

Return type
tuple[Dataset, list[str]]

Notes

- This method is responsible for trying (and possibly failing) to coerce a Dataset into a valid form with attributes and encodings. We use check_dataset_conformance to check for validation errors.

classmethod from_yaml(*product_definition_filepath: str | CloudPath | Path*)

Create a DataProductDefinition from a Libera data product definition YAML file.

Parameters
product_definition_filepath (*str | PathType*) – Path to YAML file with product and variable definitions

Returns
Configured instance with loaded metadata and optional data

Return type
DataProductDefinition

`generate_data_product_filename`(*utc_start: datetime*, *utc_end: datetime*) → *LiberaDataProductFilename*

Generate a standardized Libera data product filename.

Parameters

- **utc_start** (*datetime*) – Start time of data in the file
- **utc_end** (*datetime*) – End time of data in the file

Returns
Properly formatted filename object

Return type
LiberaDataProductFilename

`model_config: ClassVar[ConfigDict] = {'frozen': True}`

Configuration for the model, should be a dictionary conforming to [*ConfigDict*][pydantic.config.ConfigDict].

`property static_attributes`

Return product-level attributes that are statically defined (have values) in the data product definition

`class libera_utils.Manifest`(*, *manifest_type: ~libera_utils.constants.ManifestType*, *files: list[~libera_utils.io.manifest.ManifestFileRecord] = <factory>*, *configuration: dict[str, ~typing.Any] = <factory>*, *filename: ~libera_utils.io.filenaming.ManifestFilename | None = None*, *ulid_code: ~ulid.ULID | None = <factory>*)

Pydantic model for a manifest file.

Attributes

model_extra
Get extra fields set during validation.

model_fields_set
Returns the set of fields that have been explicitly set on this model instance.

Methods

add_desired_time_range(start_datetime, ...)	Add a time range to the configuration section of the manifest.
add_files(*files)	Add files to the manifest from filename
check_file_structure(file_structure, ...)	Check file structure, returning True if it is good.
copy(*[, include, exclude, update, deep])	Returns a copy of the model.
from_file(filepath)	Read a manifest file and return a Manifest object (factory method).
model_construct([_fields_set])	Creates a new instance of the *Model* class with validated data.
model_copy(*[, update, deep])	!!! abstract "Usage Documentation"
model_dump(*[, mode, include, exclude, ...])	!!! abstract "Usage Documentation"
model_dump_json(*[, indent, ensure_ascii, ...])	!!! abstract "Usage Documentation"

continues on next page

Table 3 – continued from previous page

model_json_schema([by_alias, ref_template, ...])	Generates a JSON schema for a model class.
model_parametrized_name(params)	Compute the class name for parametrizations of generic classes.
model_post_init(context, /)	Override this method to perform additional initialization after *__init__* and *model_construct*.
model_rebuild(*[, force, raise_errors, ...])	Try to rebuild the pydantic-core schema for the model.
model_validate(obj, *[, strict, extra, ...])	Validate a pydantic model instance.
model_validate_json(json_data, *[, strict, ...])	!!! abstract "Usage Documentation"
model_validate_strings(obj, *[, strict, ...])	Validate the given object with string data against the Pydantic model.
output_manifest_from_input_manifest(...)	Create Output manifest from input manifest file path, adds input files to output manifest configuration
serialize_filename(filename, _info)	Custom serializer for the manifest filename.
transform_filename(raw_filename)	Convert raw filename to ManifestFilename class if necessary.
transform_files(raw_list)	Allow for the incoming files list to have varying types.
validate_checksums()	Validate checksums of listed files
write(out_path[, filename])	Write a manifest file from a Manifest object (self).

construct
dict
from_orm
json
parse_file
parse_obj
parse_raw
schema
schema_json
update_forward_refs
validate

_generate_filename() → *ManifestFilename*

Generate a valid manifest filename

add_desired_time_range(*start_datetime: datetime*, *end_datetime: datetime*)

Add a time range to the configuration section of the manifest.

Parameters

- **start_datetime** (*datetime.datetime*) – The desired start time for the range of data in this manifest
- **end_datetime** (*datetime.datetime*) – The desired end time for the range of data in this manifest

Return type
None

add_files(**files: str | Path | S3Path*)

Add files to the manifest from filename

Parameters
files (*Union[str, Path, S3Path]*) – Path to the file to add to the manifest.

Return type
None

classmethod check_file_structure(*file_structure:* ManifestFileRecord, *existing_names: set[str]*, *existing_checksums: set[str]*) → bool

Check file structure, returning True if it is good.

classmethod from_file(*filepath: str | Path | S3Path*)

Read a manifest file and return a Manifest object (factory method).

Parameters
filepath (*Union[str, Path, S3Path]*) – Location of manifest file to read.

Returns
Pydantic model built from the json of the given manifest file.

Return type
Manifest

model_config: ClassVar[ConfigDict] = {'arbitrary_types_allowed': True}

Configuration for the model, should be a dictionary conforming to [*ConfigDict*][pydantic.config.ConfigDict].

classmethod output_manifest_from_input_manifest(*input_manifest: Path | S3Path | Manifest*) → *Manifest*

Create Output manifest from input manifest file path, adds input files to output manifest configuration

Parameters
input_manifest (*Union[Path, S3Path, 'Manifest']*) – An S3 or regular path to an input_manifest object, or the input manifest object itself

Returns
output_manifest – The newly created output manifest

Return type
Manifest

serialize_filename(*filename: str | Path | S3Path |* ManifestFilename *| None*, *_info*) → str

Custom serializer for the manifest filename.

classmethod transform_filename(*raw_filename: str |* ManifestFilename *| None*) → *ManifestFilename |* None

Convert raw filename to ManifestFilename class if necessary.

classmethod transform_files(*raw_list: list[dict | str | Path | S3Path |* ManifestFileRecord*] | None*) → list[*ManifestFileRecord*]

Allow for the incoming files list to have varying types. Convert to a standardized list of ManifestFileStructure.

validate_checksums() → None

Validate checksums of listed files

write(*out_path: str | Path | S3Path*, *filename: str = None*) → Path | S3Path

Write a manifest file from a Manifest object (self).

Parameters

- **out_path** (*Union[str, Path, S3Path]*) – Directory path to write to (directory being used loosely to refer also to an S3 bucket path).

- **filename** (*str, Optional*) – must be a valid manifest filename. If not provided, the method uses the objects internal filename attribute. If that is not set, then a filename is automatically generated.

Returns
The path where the manifest file is written.

Return type
Union[Path, S3Path]

class libera_utils.**ManifestType**(*value*)

Enumerated legal manifest type values

Methods

capitalize(/)	Return a capitalized version of the string.
casefold(/)	Return a version of the string suitable for caseless comparisons.
center(width[, fillchar])	Return a centered string of length width.
count(sub[, start[, end]])	Return the number of non-overlapping occurrences of substring sub in string S[start:end].
encode(/[, encoding, errors])	Encode the string using the codec registered for encoding.
endswith(suffix[, start[, end]])	Return True if S ends with the specified suffix, False otherwise.
expandtabs(/[, tabsize])	Return a copy where all tab characters are expanded using spaces.
find(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
format(*args, **kwargs)	Return a formatted version of S, using substitutions from args and kwargs.
format_map(mapping)	Return a formatted version of S, using substitutions from mapping.
index(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
isalnum(/)	Return True if the string is an alpha-numeric string, False otherwise.
isalpha(/)	Return True if the string is an alphabetic string, False otherwise.
isascii(/)	Return True if all characters in the string are ASCII, False otherwise.
isdecimal(/)	Return True if the string is a decimal string, False otherwise.
isdigit(/)	Return True if the string is a digit string, False otherwise.
isidentifier(/)	Return True if the string is a valid Python identifier, False otherwise.
islower(/)	Return True if the string is a lowercase string, False otherwise.
isnumeric(/)	Return True if the string is a numeric string, False otherwise.
isprintable(/)	Return True if the string is printable, False otherwise.
isspace(/)	Return True if the string is a whitespace string, False otherwise.

continues on next page

Table 4 – continued from previous page

istitle(/)	Return True if the string is a title-cased string, False otherwise.
isupper(/)	Return True if the string is an uppercase string, False otherwise.
join(iterable, /)	Concatenate any number of strings.
ljust(width[, fillchar])	Return a left-justified string of length width.
lower(/)	Return a copy of the string converted to lowercase.
lstrip([chars])	Return a copy of the string with leading whitespace removed.
maketrans(x[, y, z])	Return a translation table usable for str.translate().
partition(sep, /)	Partition the string into three parts using the given separator.
removeprefix(prefix, /)	Return a str with the given prefix string removed if present.
removesuffix(suffix, /)	Return a str with the given suffix string removed if present.
replace(old, new[, count])	Return a copy with all occurrences of substring old replaced by new.
rfind(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rindex(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rjust(width[, fillchar])	Return a right-justified string of length width.
rpartition(sep, /)	Partition the string into three parts using the given separator.
rsplit(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
rstrip([chars])	Return a copy of the string with trailing whitespace removed.
split(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
splitlines(/[, keepends])	Return a list of the lines in the string, breaking at line boundaries.
startswith(prefix[, start[, end]])	Return True if S starts with the specified prefix, False otherwise.
strip([chars])	Return a copy of the string with leading and trailing whitespace removed.
swapcase(/)	Convert uppercase characters to lowercase and lowercase characters to uppercase.
title(/)	Return a version of the string where each word is titlecased.
translate(table, /)	Replace each character in the string using the given translation table.
upper(/)	Return a copy of the string converted to uppercase.
zfill(width, /)	Pad a numeric string with zeros on the left, to fill a field of the given width.

libera_utils.**smart_copy_file**(*source_path: str | Path | S3Path*, *dest_path: str | Path | S3Path*, *delete: bool | None = False*)

Copy function that can handle local files or files in an S3 bucket. Returns the path to the newly created file as a Path or an S3Path, depending on the destination.

Parameters

- **source_path** (*Union[str, Path, S3Path]*) – Path to the source file to be copied. Files residing in an s3 bucket must begin with "s3://".
- **dest_path** (*Union[str, Path, S3Path]*) – Path to the Destination file to be copied to. Files residing in an s3 bucket must begin with "s3://".
- **delete** (*bool, optional*) – If true, deletes files copied from source (default = False)

Returns
The path to the newly created file

Return type
Path or S3Path

libera_utils.**smart_open**(*path: str | Path | S3Path*, *mode: str | None = 'rb'*, *enable_gzip: bool | None = True*)

Open function that can handle local files or files in an S3 bucket. It also correctly handles gzip files determined by a **.gz* extension.

Parameters

- **path** (*Union[str, Path, S3Path]*) – Path to the file to be opened. Files residing in an s3 bucket must begin with "s3://".
- **mode** (*str, Optional*) – Optional string specifying the mode in which the file is opened. Defaults to 'rb'.
- **enable_gzip** (*bool, Optional*) – Flag to specify that **.gz* files should be opened as a *GzipFile* object. Setting this to False is useful when creating the md5sum of a **.gz* file. Defaults to True.

Return type
IO or gzip.GzipFile

Modules

aws	
cli	Module for the Libera SDC utilities CLI
config	Configuration reader.
constants	Constants module used throughout the libera_utils package
db	db module for dyanmodb operations.
io	
kernel_maker	Module containing CLI tool for creating SPICE kernels from packets
logutil	Logging utilities
packets	Module for reading packet data using Space Packet Parser
quality_flags	Quality flag definitions
spice_utils	Modules for SPICE kernel creation, management, and usage
time	Module for dealing with time and time conventions
version	Module for anything related to package versioning

3.1.1 libera_utils.aws

Modules

constants	Constants used specifically for AWS resources and resource naming
ecr_upload	Module for uploading docker images to the ECR
processing_step_function_trigger	Module for manually triggering a step function
s3_utilities	Module for S3 cli utilities
utils	Helper functions for AWS access

libera_utils.aws.constants

Constants used specifically for AWS resources and resource naming

Classes

LiberaAccountSuffix(value)	Suffixes for the various account types
LiberaDataBucketName(value)	Names of the data archive buckets

libera_utils.aws.constants.LiberaAccountSuffix

class libera_utils.aws.constants.**LiberaAccountSuffix**(*value*)

Bases: StrEnum

Suffixes for the various account types

Methods

capitalize(/)	Return a capitalized version of the string.
casefold(/)	Return a version of the string suitable for caseless comparisons.
center(width[, fillchar])	Return a centered string of length width.
count(sub[, start[, end]])	Return the number of non-overlapping occurrences of substring sub in string S[start:end].
encode(/[, encoding, errors])	Encode the string using the codec registered for encoding.
endswith(suffix[, start[, end]])	Return True if S ends with the specified suffix, False otherwise.
expandtabs(/[, tabsize])	Return a copy where all tab characters are expanded using spaces.
find(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
format(*args, **kwargs)	Return a formatted version of S, using substitutions from args and kwargs.
format_map(mapping)	Return a formatted version of S, using substitutions from mapping.

continues on next page

Table 8 – continued from previous page

index(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
isalnum(/)	Return True if the string is an alpha-numeric string, False otherwise.
isalpha(/)	Return True if the string is an alphabetic string, False otherwise.
isascii(/)	Return True if all characters in the string are ASCII, False otherwise.
isdecimal(/)	Return True if the string is a decimal string, False otherwise.
isdigit(/)	Return True if the string is a digit string, False otherwise.
isidentifier(/)	Return True if the string is a valid Python identifier, False otherwise.
islower(/)	Return True if the string is a lowercase string, False otherwise.
isnumeric(/)	Return True if the string is a numeric string, False otherwise.
isprintable(/)	Return True if the string is printable, False otherwise.
isspace(/)	Return True if the string is a whitespace string, False otherwise.
istitle(/)	Return True if the string is a title-cased string, False otherwise.
isupper(/)	Return True if the string is an uppercase string, False otherwise.
join(iterable, /)	Concatenate any number of strings.
ljust(width[, fillchar])	Return a left-justified string of length width.
lower(/)	Return a copy of the string converted to lowercase.
lstrip([chars])	Return a copy of the string with leading whitespace removed.
maketrans(x[, y, z])	Return a translation table usable for str.translate().
partition(sep, /)	Partition the string into three parts using the given separator.
removeprefix(prefix, /)	Return a str with the given prefix string removed if present.
removesuffix(suffix, /)	Return a str with the given suffix string removed if present.
replace(old, new[, count])	Return a copy with all occurrences of substring old replaced by new.
rfind(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rindex(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rjust(width[, fillchar])	Return a right-justified string of length width.
rpartition(sep, /)	Partition the string into three parts using the given separator.
rsplit(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
rstrip([chars])	Return a copy of the string with trailing whitespace removed.
split(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.

continues on next page

Table 8 – continued from previous page

splitlines(/[, keepends])	Return a list of the lines in the string, breaking at line boundaries.
startswith(prefix[, start[, end]])	Return True if S starts with the specified prefix, False otherwise.
strip([chars])	Return a copy of the string with leading and trailing whitespace removed.
swapcase(/)	Convert uppercase characters to lowercase and lowercase characters to uppercase.
title(/)	Return a version of the string where each word is titlecased.
translate(table, /)	Replace each character in the string using the given translation table.
upper(/)	Return a copy of the string converted to uppercase.
zfill(width, /)	Pad a numeric string with zeros on the left, to fill a field of the given width.

__init__(**args*, ***kwds*)

Methods

encode(/[, encoding, errors])	Encode the string using the codec registered for encoding.
replace(old, new[, count])	Return a copy with all occurrences of substring old replaced by new.
split(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
rsplit(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
join(iterable, /)	Concatenate any number of strings.
capitalize(/)	Return a capitalized version of the string.
casefold(/)	Return a version of the string suitable for caseless comparisons.
title(/)	Return a version of the string where each word is titlecased.
center(width[, fillchar])	Return a centered string of length width.
count(sub[, start[, end]])	Return the number of non-overlapping occurrences of substring sub in string S[start:end].
expandtabs(/[, tabsize])	Return a copy where all tab characters are expanded using spaces.
find(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
partition(sep, /)	Partition the string into three parts using the given separator.
index(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
ljust(width[, fillchar])	Return a left-justified string of length width.
lower(/)	Return a copy of the string converted to lowercase.
lstrip([chars])	Return a copy of the string with leading whitespace removed.

continues on next page

Table 9 – continued from previous page

rfind(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rindex(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rjust(width[, fillchar])	Return a right-justified string of length width.
rstrip([chars])	Return a copy of the string with trailing whitespace removed.
rpartition(sep, /)	Partition the string into three parts using the given separator.
splitlines(/[, keepends])	Return a list of the lines in the string, breaking at line boundaries.
strip([chars])	Return a copy of the string with leading and trailing whitespace removed.
swapcase(/)	Convert uppercase characters to lowercase and lowercase characters to uppercase.
translate(table, /)	Replace each character in the string using the given translation table.
upper(/)	Return a copy of the string converted to uppercase.
startswith(prefix[, start[, end]])	Return True if S starts with the specified prefix, False otherwise.
endswith(suffix[, start[, end]])	Return True if S ends with the specified suffix, False otherwise.
removeprefix(prefix, /)	Return a str with the given prefix string removed if present.
removesuffix(suffix, /)	Return a str with the given suffix string removed if present.
isascii(/)	Return True if all characters in the string are ASCII, False otherwise.
islower(/)	Return True if the string is a lowercase string, False otherwise.
isupper(/)	Return True if the string is an uppercase string, False otherwise.
istitle(/)	Return True if the string is a title-cased string, False otherwise.
isspace(/)	Return True if the string is a whitespace string, False otherwise.
isdecimal(/)	Return True if the string is a decimal string, False otherwise.
isdigit(/)	Return True if the string is a digit string, False otherwise.
isnumeric(/)	Return True if the string is a numeric string, False otherwise.
isalpha(/)	Return True if the string is an alphabetic string, False otherwise.
isalnum(/)	Return True if the string is an alpha-numeric string, False otherwise.
isidentifier(/)	Return True if the string is a valid Python identifier, False otherwise.
isprintable(/)	Return True if the string is printable, False otherwise.
zfill(width, /)	Pad a numeric string with zeros on the left, to fill a field of the given width.

continues on next page

Table 9 – continued from previous page

format(*args, **kwargs)	Return a formatted version of S, using substitutions from args and kwargs.
format_map(mapping)	Return a formatted version of S, using substitutions from mapping.
maketrans(x[, y, z])	Return a translation table usable for str.translate().

Attributes

STAGE
PROD
DEV

capitalize(*/*)

Return a capitalized version of the string.

More specifically, make the first character have upper case and the rest lower case.

casefold(*/*)

Return a version of the string suitable for caseless comparisons.

center(*width*, *fillchar=' '*, */ (Positional-only parameter separator (PEP 570))*)

Return a centered string of length width.

Padding is done using the specified fill character (default is a space).

count(*sub*[, *start*[, *end*]]) → int

Return the number of non-overlapping occurrences of substring sub in string S[start:end]. Optional arguments start and end are interpreted as in slice notation.

encode(*/*, *encoding='utf-8'*, *errors='strict'*)

Encode the string using the codec registered for encoding.

encoding
The encoding in which to encode the string.

errors
The error handling scheme to use for encoding errors. The default is 'strict' meaning that encoding errors raise a UnicodeEncodeError. Other possible values are 'ignore', 'replace' and 'xmlcharrefreplace' as well as any other name registered with codecs.register_error that can handle UnicodeEncodeErrors.

endswith(*suffix*[, *start*[, *end*]]) → bool

Return True if S ends with the specified suffix, False otherwise. With optional start, test S beginning at that position. With optional end, stop comparing S at that position. suffix can also be a tuple of strings to try.

expandtabs(*/*, *tabsize=8*)

Return a copy where all tab characters are expanded using spaces.

If tabsize is not given, a tab size of 8 characters is assumed.

find(*sub*[, *start*[, *end*]]) → int

Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end]. Optional arguments start and end are interpreted as in slice notation.

Return -1 on failure.

format(**args*, ***kwargs*) → str

Return a formatted version of S, using substitutions from args and kwargs. The substitutions are identified by braces ('{' and '}').

format_map(*mapping*) → str

Return a formatted version of S, using substitutions from mapping. The substitutions are identified by braces ('{' and '}').

index(*sub*[, *start*[, *end*]]) → int

Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end]. Optional arguments start and end are interpreted as in slice notation.

Raises ValueError when the substring is not found.

isalnum(/)

Return True if the string is an alpha-numeric string, False otherwise.

A string is alpha-numeric if all characters in the string are alpha-numeric and there is at least one character in the string.

isalpha(/)

Return True if the string is an alphabetic string, False otherwise.

A string is alphabetic if all characters in the string are alphabetic and there is at least one character in the string.

isascii(/)

Return True if all characters in the string are ASCII, False otherwise.

ASCII characters have code points in the range U+0000-U+007F. Empty string is ASCII too.

isdecimal(/)

Return True if the string is a decimal string, False otherwise.

A string is a decimal string if all characters in the string are decimal and there is at least one character in the string.

isdigit(/)

Return True if the string is a digit string, False otherwise.

A string is a digit string if all characters in the string are digits and there is at least one character in the string.

isidentifier(/)

Return True if the string is a valid Python identifier, False otherwise.

Call keyword.iskeyword(s) to test whether string s is a reserved identifier, such as "def" or "class".

islower(/)

Return True if the string is a lowercase string, False otherwise.

A string is lowercase if all cased characters in the string are lowercase and there is at least one cased character in the string.

isnumeric(/)

Return True if the string is a numeric string, False otherwise.

A string is numeric if all characters in the string are numeric and there is at least one character in the string.

isprintable(*/*)

Return True if the string is printable, False otherwise.

A string is printable if all of its characters are considered printable in repr() or if it is empty.

isspace(*/*)

Return True if the string is a whitespace string, False otherwise.

A string is whitespace if all characters in the string are whitespace and there is at least one character in the string.

istitle(*/*)

Return True if the string is a title-cased string, False otherwise.

In a title-cased string, upper- and title-case characters may only follow uncased characters and lowercase characters only cased ones.

isupper(*/*)

Return True if the string is an uppercase string, False otherwise.

A string is uppercase if all cased characters in the string are uppercase and there is at least one cased character in the string.

join(*iterable*, */*)

Concatenate any number of strings.

The string whose method is called is inserted in between each given string. The result is returned as a new string.

Example: '.'.join(['ab', 'pq', 'rs']) -> 'ab.pq.rs'

ljust(*width*, *fillchar=' '*, */*)

Return a left-justified string of length width.

Padding is done using the specified fill character (default is a space).

lower(*/*)

Return a copy of the string converted to lowercase.

lstrip(*chars=None*, */*)

Return a copy of the string with leading whitespace removed.

If chars is given and not None, remove characters in chars instead.

static maketrans(*x*, *y=<unrepresentable>*, *z=<unrepresentable>*, */*)

Return a translation table usable for str.translate().

If there is only one argument, it must be a dictionary mapping Unicode ordinals (integers) or characters to Unicode ordinals, strings or None. Character keys will be then converted to ordinals. If there are two arguments, they must be strings of equal length, and in the resulting dictionary, each character in x will be mapped to the character at the same position in y. If there is a third argument, it must be a string, whose characters will be mapped to None in the result.

partition(*sep*, */*)

Partition the string into three parts using the given separator.

This will search for the separator in the string. If the separator is found, returns a 3-tuple containing the part before the separator, the separator itself, and the part after it.

If the separator is not found, returns a 3-tuple containing the original string and two empty strings.

removeprefix(*prefix*, /)

Return a str with the given prefix string removed if present.

If the string starts with the prefix string, return string[len(prefix):]. Otherwise, return a copy of the original string.

removesuffix(*suffix*, /)

Return a str with the given suffix string removed if present.

If the string ends with the suffix string and that suffix is not empty, return string[:-len(suffix)]. Otherwise, return a copy of the original string.

replace(*old*, *new*, *count=-1*, /)

Return a copy with all occurrences of substring old replaced by new.

count
Maximum number of occurrences to replace. -1 (the default value) means replace all occurrences.

If the optional argument count is given, only the first count occurrences are replaced.

rfind(*sub*[, *start*[, *end*]]) → int

Return the highest index in S where substring sub is found, such that sub is contained within S[start:end]. Optional arguments start and end are interpreted as in slice notation.

Return -1 on failure.

rindex(*sub*[, *start*[, *end*]]) → int

Return the highest index in S where substring sub is found, such that sub is contained within S[start:end]. Optional arguments start and end are interpreted as in slice notation.

Raises ValueError when the substring is not found.

rjust(*width*, *fillchar=' '*, /)

Return a right-justified string of length width.

Padding is done using the specified fill character (default is a space).

rpartition(*sep*, /)

Partition the string into three parts using the given separator.

This will search for the separator in the string, starting at the end. If the separator is found, returns a 3-tuple containing the part before the separator, the separator itself, and the part after it.

If the separator is not found, returns a 3-tuple containing two empty strings and the original string.

rsplit(*/*, *sep=None*, *maxsplit=-1*)

Return a list of the substrings in the string, using sep as the separator string.

sep
The separator used to split the string.

When set to None (the default value), will split on any whitespace character (including n r t f and spaces) and will discard empty strings from the result.

maxsplit
Maximum number of splits. -1 (the default value) means no limit.

Splitting starts at the end of the string and works to the front.

rstrip(*chars=None*, */*)

Return a copy of the string with trailing whitespace removed.

If chars is given and not None, remove characters in chars instead.

split(*/*, *sep=None*, *maxsplit=-1*)

Return a list of the substrings in the string, using sep as the separator string.

sep
The separator used to split the string.

When set to None (the default value), will split on any whitespace character (including n r t f and spaces) and will discard empty strings from the result.

maxsplit
Maximum number of splits. -1 (the default value) means no limit.

Splitting starts at the front of the string and works to the end.

Note, str.split() is mainly useful for data that has been intentionally delimited. With natural text that includes punctuation, consider using the regular expression module.

splitlines(*/*, *keepends=False*)

Return a list of the lines in the string, breaking at line boundaries.

Line breaks are not included in the resulting list unless keepends is given and true.

startswith(*prefix*[, *start*[, *end*]]) → bool

Return True if S starts with the specified prefix, False otherwise. With optional start, test S beginning at that position. With optional end, stop comparing S at that position. prefix can also be a tuple of strings to try.

strip(*chars=None*, */*)

Return a copy of the string with leading and trailing whitespace removed.

If chars is given and not None, remove characters in chars instead.

swapcase(*/*)

Convert uppercase characters to lowercase and lowercase characters to uppercase.

title(*/*)

Return a version of the string where each word is titlecased.

More specifically, words start with uppercased characters and all remaining cased characters have lower case.

translate(*table*, */*)

Replace each character in the string using the given translation table.

table
Translation table, which must be a mapping of Unicode ordinals to Unicode ordinals, strings, or None.

The table must implement lookup/indexing via __getitem__, for instance a dictionary or list. If this operation raises LookupError, the character is left untouched. Characters mapped to None are deleted.

upper(*/*)

Return a copy of the string converted to uppercase.

zfill(*width*, */*)

Pad a numeric string with zeros on the left, to fill a field of the given width.

The string is never truncated.

libera_utils.aws.constants.LiberaDataBucketName

class libera_utils.aws.constants.**LiberaDataBucketName**(*value*)

Bases: StrEnum

Names of the data archive buckets

Methods

capitalize(/)	Return a capitalized version of the string.
casefold(/)	Return a version of the string suitable for caseless comparisons.
center(width[, fillchar])	Return a centered string of length width.
count(sub[, start[, end]])	Return the number of non-overlapping occurrences of substring sub in string S[start:end].
encode(/[, encoding, errors])	Encode the string using the codec registered for encoding.
endswith(suffix[, start[, end]])	Return True if S ends with the specified suffix, False otherwise.
expandtabs(/[, tabsize])	Return a copy where all tab characters are expanded using spaces.
find(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
format(*args, **kwargs)	Return a formatted version of S, using substitutions from args and kwargs.
format_map(mapping)	Return a formatted version of S, using substitutions from mapping.
index(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
isalnum(/)	Return True if the string is an alpha-numeric string, False otherwise.
isalpha(/)	Return True if the string is an alphabetic string, False otherwise.
isascii(/)	Return True if all characters in the string are ASCII, False otherwise.
isdecimal(/)	Return True if the string is a decimal string, False otherwise.
isdigit(/)	Return True if the string is a digit string, False otherwise.
isidentifier(/)	Return True if the string is a valid Python identifier, False otherwise.
islower(/)	Return True if the string is a lowercase string, False otherwise.
isnumeric(/)	Return True if the string is a numeric string, False otherwise.
isprintable(/)	Return True if the string is printable, False otherwise.
isspace(/)	Return True if the string is a whitespace string, False otherwise.
istitle(/)	Return True if the string is a title-cased string, False otherwise.
isupper(/)	Return True if the string is an uppercase string, False otherwise.
join(iterable, /)	Concatenate any number of strings.

continues on next page

Table 11 – continued from previous page

ljust(width[, fillchar])	Return a left-justified string of length width.
lower(/)	Return a copy of the string converted to lowercase.
lstrip([chars])	Return a copy of the string with leading whitespace removed.
maketrans(x[, y, z])	Return a translation table usable for str.translate().
partition(sep, /)	Partition the string into three parts using the given separator.
removeprefix(prefix, /)	Return a str with the given prefix string removed if present.
removesuffix(suffix, /)	Return a str with the given suffix string removed if present.
replace(old, new[, count])	Return a copy with all occurrences of substring old replaced by new.
rfind(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rindex(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rjust(width[, fillchar])	Return a right-justified string of length width.
rpartition(sep, /)	Partition the string into three parts using the given separator.
rsplit(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
rstrip([chars])	Return a copy of the string with trailing whitespace removed.
split(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
splitlines(/[, keepends])	Return a list of the lines in the string, breaking at line boundaries.
startswith(prefix[, start[, end]])	Return True if S starts with the specified prefix, False otherwise.
strip([chars])	Return a copy of the string with leading and trailing whitespace removed.
swapcase(/)	Convert uppercase characters to lowercase and lowercase characters to uppercase.
title(/)	Return a version of the string where each word is titlecased.
translate(table, /)	Replace each character in the string using the given translation table.
upper(/)	Return a copy of the string converted to uppercase.
zfill(width, /)	Pad a numeric string with zeros on the left, to fill a field of the given width.

__init__(**args*, ***kwds*)

Methods

encode(/[, encoding, errors])	Encode the string using the codec registered for encoding.
replace(old, new[, count])	Return a copy with all occurrences of substring old replaced by new.

continues on next page

Table 12 – continued from previous page

split(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
rsplit(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
join(iterable, /)	Concatenate any number of strings.
capitalize(/)	Return a capitalized version of the string.
casefold(/)	Return a version of the string suitable for caseless comparisons.
title(/)	Return a version of the string where each word is titlecased.
center(width[, fillchar])	Return a centered string of length width.
count(sub[, start[, end]])	Return the number of non-overlapping occurrences of substring sub in string S[start:end].
expandtabs(/[, tabsize])	Return a copy where all tab characters are expanded using spaces.
find(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
partition(sep, /)	Partition the string into three parts using the given separator.
index(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
ljust(width[, fillchar])	Return a left-justified string of length width.
lower(/)	Return a copy of the string converted to lowercase.
lstrip([chars])	Return a copy of the string with leading whitespace removed.
rfind(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rindex(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rjust(width[, fillchar])	Return a right-justified string of length width.
rstrip([chars])	Return a copy of the string with trailing whitespace removed.
rpartition(sep, /)	Partition the string into three parts using the given separator.
splitlines(/[, keepends])	Return a list of the lines in the string, breaking at line boundaries.
strip([chars])	Return a copy of the string with leading and trailing whitespace removed.
swapcase(/)	Convert uppercase characters to lowercase and lowercase characters to uppercase.
translate(table, /)	Replace each character in the string using the given translation table.
upper(/)	Return a copy of the string converted to uppercase.
startswith(prefix[, start[, end]])	Return True if S starts with the specified prefix, False otherwise.
endswith(suffix[, start[, end]])	Return True if S ends with the specified suffix, False otherwise.
removeprefix(prefix, /)	Return a str with the given prefix string removed if present.
removesuffix(suffix, /)	Return a str with the given suffix string removed if present.

continues on next page

Table 12 – continued from previous page

isascii(/)	Return True if all characters in the string are ASCII, False otherwise.
islower(/)	Return True if the string is a lowercase string, False otherwise.
isupper(/)	Return True if the string is an uppercase string, False otherwise.
istitle(/)	Return True if the string is a title-cased string, False otherwise.
isspace(/)	Return True if the string is a whitespace string, False otherwise.
isdecimal(/)	Return True if the string is a decimal string, False otherwise.
isdigit(/)	Return True if the string is a digit string, False otherwise.
isnumeric(/)	Return True if the string is a numeric string, False otherwise.
isalpha(/)	Return True if the string is an alphabetic string, False otherwise.
isalnum(/)	Return True if the string is an alpha-numeric string, False otherwise.
isidentifier(/)	Return True if the string is a valid Python identifier, False otherwise.
isprintable(/)	Return True if the string is printable, False otherwise.
zfill(width, /)	Pad a numeric string with zeros on the left, to fill a field of the given width.
format(*args, **kwargs)	Return a formatted version of S, using substitutions from args and kwargs.
format_map(mapping)	Return a formatted version of S, using substitutions from mapping.
maketrans(x[, y, z])	Return a translation table usable for str.translate().

Attributes

L0_ARCHIVE_BUCKET
L1A_ARCHIVE_BUCKET
SPICE_ARCHIVE_BUCKET
ANCILLARY_ARCHIVE_BUCKET
L1B_ARCHIVE_BUCKET
L2_ARCHIVE_BUCKET

capitalize(/)

Return a capitalized version of the string.

More specifically, make the first character have upper case and the rest lower case.

casefold(/)

Return a version of the string suitable for caseless comparisons.

center(*width*, *fillchar=' '*, /)

Return a centered string of length width.

Padding is done using the specified fill character (default is a space).

count(*sub*[, *start*[, *end*]]) → int

Return the number of non-overlapping occurrences of substring sub in string S[start:end]. Optional arguments start and end are interpreted as in slice notation.

encode(/, *encoding='utf-8'*, *errors='strict'*)

Encode the string using the codec registered for encoding.

encoding
The encoding in which to encode the string.

errors
The error handling scheme to use for encoding errors. The default is 'strict' meaning that encoding errors raise a UnicodeEncodeError. Other possible values are 'ignore', 'replace' and 'xmlcharrefreplace' as well as any other name registered with codecs.register_error that can handle UnicodeEncodeErrors.

endswith(*suffix*[, *start*[, *end*]]) → bool

Return True if S ends with the specified suffix, False otherwise. With optional start, test S beginning at that position. With optional end, stop comparing S at that position. suffix can also be a tuple of strings to try.

expandtabs(/, *tabsize=8*)

Return a copy where all tab characters are expanded using spaces.

If tabsize is not given, a tab size of 8 characters is assumed.

find(*sub*[, *start*[, *end*]]) → int

Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end]. Optional arguments start and end are interpreted as in slice notation.

Return -1 on failure.

format(**args*, ***kwargs*) → str

Return a formatted version of S, using substitutions from args and kwargs. The substitutions are identified by braces ('{' and '}').

format_map(*mapping*) → str

Return a formatted version of S, using substitutions from mapping. The substitutions are identified by braces ('{' and '}').

index(*sub*[, *start*[, *end*]]) → int

Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end]. Optional arguments start and end are interpreted as in slice notation.

Raises ValueError when the substring is not found.

isalnum(/)

Return True if the string is an alpha-numeric string, False otherwise.

A string is alpha-numeric if all characters in the string are alpha-numeric and there is at least one character in the string.

isalpha(/)

Return True if the string is an alphabetic string, False otherwise.

A string is alphabetic if all characters in the string are alphabetic and there is at least one character in the string.

isascii(/)

Return True if all characters in the string are ASCII, False otherwise.

ASCII characters have code points in the range U+0000-U+007F. Empty string is ASCII too.

isdecimal(/)

Return True if the string is a decimal string, False otherwise.

A string is a decimal string if all characters in the string are decimal and there is at least one character in the string.

isdigit(/)

Return True if the string is a digit string, False otherwise.

A string is a digit string if all characters in the string are digits and there is at least one character in the string.

isidentifier(/)

Return True if the string is a valid Python identifier, False otherwise.

Call keyword.iskeyword(s) to test whether string s is a reserved identifier, such as “def” or “class”.

islower(/)

Return True if the string is a lowercase string, False otherwise.

A string is lowercase if all cased characters in the string are lowercase and there is at least one cased character in the string.

isnumeric(/)

Return True if the string is a numeric string, False otherwise.

A string is numeric if all characters in the string are numeric and there is at least one character in the string.

isprintable(/)

Return True if the string is printable, False otherwise.

A string is printable if all of its characters are considered printable in repr() or if it is empty.

isspace(/)

Return True if the string is a whitespace string, False otherwise.

A string is whitespace if all characters in the string are whitespace and there is at least one character in the string.

istitle(/)

Return True if the string is a title-cased string, False otherwise.

In a title-cased string, upper- and title-case characters may only follow uncased characters and lowercase characters only cased ones.

isupper(/)

Return True if the string is an uppercase string, False otherwise.

A string is uppercase if all cased characters in the string are uppercase and there is at least one cased character in the string.

join(*iterable*, /)

Concatenate any number of strings.

The string whose method is called is inserted in between each given string. The result is returned as a new string.

Example: '.'.join(['ab', 'pq', 'rs']) -> 'ab.pq.rs'

ljust(*width*, *fillchar=' '*, /)

Return a left-justified string of length width.

Padding is done using the specified fill character (default is a space).

lower(/)

Return a copy of the string converted to lowercase.

lstrip(*chars=None*, /)

Return a copy of the string with leading whitespace removed.

If chars is given and not None, remove characters in chars instead.

static maketrans(*x*, *y=<unrepresentable>*, *z=<unrepresentable>*, /)

Return a translation table usable for str.translate().

If there is only one argument, it must be a dictionary mapping Unicode ordinals (integers) or characters to Unicode ordinals, strings or None. Character keys will be then converted to ordinals. If there are two arguments, they must be strings of equal length, and in the resulting dictionary, each character in x will be mapped to the character at the same position in y. If there is a third argument, it must be a string, whose characters will be mapped to None in the result.

partition(*sep*, /)

Partition the string into three parts using the given separator.

This will search for the separator in the string. If the separator is found, returns a 3-tuple containing the part before the separator, the separator itself, and the part after it.

If the separator is not found, returns a 3-tuple containing the original string and two empty strings.

removeprefix(*prefix*, /)

Return a str with the given prefix string removed if present.

If the string starts with the prefix string, return string[len(prefix):]. Otherwise, return a copy of the original string.

removesuffix(*suffix*, /)

Return a str with the given suffix string removed if present.

If the string ends with the suffix string and that suffix is not empty, return string[:-len(suffix)]. Otherwise, return a copy of the original string.

replace(*old*, *new*, *count=-1*, /)

Return a copy with all occurrences of substring old replaced by new.

count
Maximum number of occurrences to replace. -1 (the default value) means replace all occurrences.

If the optional argument count is given, only the first count occurrences are replaced.

rfind(*sub*[, *start*[, *end*]]) → int

Return the highest index in S where substring sub is found, such that sub is contained within S[start:end]. Optional arguments start and end are interpreted as in slice notation.

Return -1 on failure.

rindex(*sub*[, *start*[, *end*]]) → int

Return the highest index in S where substring sub is found, such that sub is contained within S[start:end]. Optional arguments start and end are interpreted as in slice notation.

Raises ValueError when the substring is not found.

rjust(*width*, *fillchar=' '*, */*)

Return a right-justified string of length width.

Padding is done using the specified fill character (default is a space).

rpartition(*sep*, */*)

Partition the string into three parts using the given separator.

This will search for the separator in the string, starting at the end. If the separator is found, returns a 3-tuple containing the part before the separator, the separator itself, and the part after it.

If the separator is not found, returns a 3-tuple containing two empty strings and the original string.

rsplit(*/*, *sep=None*, *maxsplit=-1*)

Return a list of the substrings in the string, using sep as the separator string.

sep
The separator used to split the string.

When set to None (the default value), will split on any whitespace character (including n r t f and spaces) and will discard empty strings from the result.

maxsplit
Maximum number of splits. -1 (the default value) means no limit.

Splitting starts at the end of the string and works to the front.

rstrip(*chars=None*, */*)

Return a copy of the string with trailing whitespace removed.

If chars is given and not None, remove characters in chars instead.

split(*/*, *sep=None*, *maxsplit=-1*)

Return a list of the substrings in the string, using sep as the separator string.

sep
The separator used to split the string.

When set to None (the default value), will split on any whitespace character (including n r t f and spaces) and will discard empty strings from the result.

maxsplit
Maximum number of splits. -1 (the default value) means no limit.

Splitting starts at the front of the string and works to the end.

Note, str.split() is mainly useful for data that has been intentionally delimited. With natural text that includes punctuation, consider using the regular expression module.

splitlines(*/*, *keepends=False*)

Return a list of the lines in the string, breaking at line boundaries.

Line breaks are not included in the resulting list unless keepends is given and true.

startswith(*prefix*[, *start*[, *end*]]) → bool

Return True if S starts with the specified prefix, False otherwise. With optional start, test S beginning at that position. With optional end, stop comparing S at that position. prefix can also be a tuple of strings to try.

strip(*chars=None*, */*)

Return a copy of the string with leading and trailing whitespace removed.

If chars is given and not None, remove characters in chars instead.

swapcase(*/*)

Convert uppercase characters to lowercase and lowercase characters to uppercase.

title(*/*)

Return a version of the string where each word is titlecased.

More specifically, words start with uppercased characters and all remaining cased characters have lower case.

translate(*table*, */*)

Replace each character in the string using the given translation table.

table
Translation table, which must be a mapping of Unicode ordinals to Unicode ordinals, strings, or None.

The table must implement lookup/indexing via __getitem__, for instance a dictionary or list. If this operation raises LookupError, the character is left untouched. Characters mapped to None are deleted.

upper(*/*)

Return a copy of the string converted to uppercase.

zfill(*width*, */*)

Pad a numeric string with zeros on the left, to fill a field of the given width.

The string is never truncated.

class libera_utils.aws.constants.**LiberaAccountSuffix**(*value*)

Suffixes for the various account types

Methods

capitalize(/)	Return a capitalized version of the string.
casefold(/)	Return a version of the string suitable for caseless comparisons.
center(width[, fillchar])	Return a centered string of length width.
count(sub[, start[, end]])	Return the number of non-overlapping occurrences of substring sub in string S[start:end].
encode(/[, encoding, errors])	Encode the string using the codec registered for encoding.
endswith(suffix[, start[, end]])	Return True if S ends with the specified suffix, False otherwise.

continues on next page

Table 14 – continued from previous page

expandtabs(/[, tabsize])	Return a copy where all tab characters are expanded using spaces.
find(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
format(*args, **kwargs)	Return a formatted version of S, using substitutions from args and kwargs.
format_map(mapping)	Return a formatted version of S, using substitutions from mapping.
index(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
isalnum(/)	Return True if the string is an alpha-numeric string, False otherwise.
isalpha(/)	Return True if the string is an alphabetic string, False otherwise.
isascii(/)	Return True if all characters in the string are ASCII, False otherwise.
isdecimal(/)	Return True if the string is a decimal string, False otherwise.
isdigit(/)	Return True if the string is a digit string, False otherwise.
isidentifier(/)	Return True if the string is a valid Python identifier, False otherwise.
islower(/)	Return True if the string is a lowercase string, False otherwise.
isnumeric(/)	Return True if the string is a numeric string, False otherwise.
isprintable(/)	Return True if the string is printable, False otherwise.
isspace(/)	Return True if the string is a whitespace string, False otherwise.
istitle(/)	Return True if the string is a title-cased string, False otherwise.
isupper(/)	Return True if the string is an uppercase string, False otherwise.
join(iterable, /)	Concatenate any number of strings.
ljust(width[, fillchar])	Return a left-justified string of length width.
lower(/)	Return a copy of the string converted to lowercase.
lstrip([chars])	Return a copy of the string with leading whitespace removed.
maketrans(x[, y, z])	Return a translation table usable for str.translate().
partition(sep, /)	Partition the string into three parts using the given separator.
removeprefix(prefix, /)	Return a str with the given prefix string removed if present.
removesuffix(suffix, /)	Return a str with the given suffix string removed if present.
replace(old, new[, count])	Return a copy with all occurrences of substring old replaced by new.
rfind(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rindex(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rjust(width[, fillchar])	Return a right-justified string of length width.

continues on next page

Table 14 – continued from previous page

rpartition(sep, /)	Partition the string into three parts using the given separator.
rsplit(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
rstrip([chars])	Return a copy of the string with trailing whitespace removed.
split(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
splitlines(/[, keepends])	Return a list of the lines in the string, breaking at line boundaries.
startswith(prefix[, start[, end]])	Return True if S starts with the specified prefix, False otherwise.
strip([chars])	Return a copy of the string with leading and trailing whitespace removed.
swapcase(/)	Convert uppercase characters to lowercase and lowercase characters to uppercase.
title(/)	Return a version of the string where each word is titlecased.
translate(table, /)	Replace each character in the string using the given translation table.
upper(/)	Return a copy of the string converted to uppercase.
zfill(width, /)	Pad a numeric string with zeros on the left, to fill a field of the given width.

class libera_utils.aws.constants.**LiberaDataBucketName**(*value*)

Names of the data archive buckets

Methods

capitalize(/)	Return a capitalized version of the string.
casefold(/)	Return a version of the string suitable for caseless comparisons.
center(width[, fillchar])	Return a centered string of length width.
count(sub[, start[, end]])	Return the number of non-overlapping occurrences of substring sub in string S[start:end].
encode(/[, encoding, errors])	Encode the string using the codec registered for encoding.
endswith(suffix[, start[, end]])	Return True if S ends with the specified suffix, False otherwise.
expandtabs(/[, tabsize])	Return a copy where all tab characters are expanded using spaces.
find(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
format(*args, **kwargs)	Return a formatted version of S, using substitutions from args and kwargs.
format_map(mapping)	Return a formatted version of S, using substitutions from mapping.
index(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].

continues on next page

Table 15 – continued from previous page

isalnum(/)	Return True if the string is an alpha-numeric string, False otherwise.
isalpha(/)	Return True if the string is an alphabetic string, False otherwise.
isascii(/)	Return True if all characters in the string are ASCII, False otherwise.
isdecimal(/)	Return True if the string is a decimal string, False otherwise.
isdigit(/)	Return True if the string is a digit string, False otherwise.
isidentifier(/)	Return True if the string is a valid Python identifier, False otherwise.
islower(/)	Return True if the string is a lowercase string, False otherwise.
isnumeric(/)	Return True if the string is a numeric string, False otherwise.
isprintable(/)	Return True if the string is printable, False otherwise.
isspace(/)	Return True if the string is a whitespace string, False otherwise.
istitle(/)	Return True if the string is a title-cased string, False otherwise.
isupper(/)	Return True if the string is an uppercase string, False otherwise.
join(iterable, /)	Concatenate any number of strings.
ljust(width[, fillchar])	Return a left-justified string of length width.
lower(/)	Return a copy of the string converted to lowercase.
lstrip([chars])	Return a copy of the string with leading whitespace removed.
maketrans(x[, y, z])	Return a translation table usable for str.translate().
partition(sep, /)	Partition the string into three parts using the given separator.
removeprefix(prefix, /)	Return a str with the given prefix string removed if present.
removesuffix(suffix, /)	Return a str with the given suffix string removed if present.
replace(old, new[, count])	Return a copy with all occurrences of substring old replaced by new.
rfind(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rindex(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rjust(width[, fillchar])	Return a right-justified string of length width.
rpartition(sep, /)	Partition the string into three parts using the given separator.
rsplit(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
rstrip([chars])	Return a copy of the string with trailing whitespace removed.
split(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
splitlines(/[, keepends])	Return a list of the lines in the string, breaking at line boundaries.

continues on next page

Table 15 – continued from previous page

startswith(prefix[, start[, end]])	Return True if S starts with the specified prefix, False otherwise.
strip([chars])	Return a copy of the string with leading and trailing whitespace removed.
swapcase(/)	Convert uppercase characters to lowercase and lowercase characters to uppercase.
title(/)	Return a version of the string where each word is titlecased.
translate(table, /)	Replace each character in the string using the given translation table.
upper(/)	Return a copy of the string converted to uppercase.
zfill(width, /)	Pad a numeric string with zeros on the left, to fill a field of the given width.

libera_utils.aws.ecr_upload

Module for uploading docker images to the ECR

Functions

build_docker_image(context_dir, image_name)	Build a Docker image from a specified directory and tag it with a custom name.
ecr_upload_cli_handler(parsed_args)	CLI handler function for ecr-upload CLI subcommand.
push_image_to_ecr(image_name, image_tag, ...)	Push a Docker image to Amazon ECR with robust authentication handling.

libera_utils.aws.ecr_upload.build_docker_image

libera_utils.aws.ecr_upload.**build_docker_image**(*context_dir: str | Path*, *image_name: str*, *tag: str = 'latest'*, *target: str | None = None*, *platform: str = 'linux/amd64'*) → None

Build a Docker image from a specified directory and tag it with a custom name.

Parameters

- **context_dir** (*Union[str, Path]*) – The path to the directory containing the Dockerfile and other build context.
- **image_name** (*str*) – The name to give the Docker image.
- **tag** (*str, optional*) – The tag to apply to the image (default is 'latest').
- **target** (*Optional[str]*) – Name of the target to build.
- **platform** (*str*) – Default "linux/amd64".

Raises

ValueError – If the specified directory does not exist or the build fails.

libera_utils.aws.ecr_upload.ecr_upload_cli_handler

libera_utils.aws.ecr_upload.**ecr_upload_cli_handler**(*parsed_args: Namespace*) → None

CLI handler function for ecr-upload CLI subcommand.

Parameters
parsed_args (*argparse.Namespace*) – Namespace of parsed CLI arguments

Return type
None

libera_utils.aws.ecr_upload.push_image_to_ecr

libera_utils.aws.ecr_upload.**push_image_to_ecr**(*image_name: str*, *image_tag: str*, *processing_step_id: str | ProcessingStepIdentifier*, ***, *ecr_image_tags: list[str] = None*, *region_name: str = 'us-west-2'*, *ignore_docker_config: bool = False*, *max_retries: int = 1*) → None

Push a Docker image to Amazon ECR with robust authentication handling.

This function handles ECR authentication by obtaining fresh credentials for each push operation, preventing authentication token expiration issues during multi-tag pushes.

Parameters

- **image_name** (*str*) – Local name of the Docker image
- **image_tag** (*str*) – Local tag of the Docker image (often 'latest')
- **processing_step_id** (*Union[str, ProcessingStepIdentifier]*) – Processing step ID string or object used to determine ECR repository name. L0 processing step IDs are not supported as they have no associated ECR.
- **ecr_image_tags** (*Optional[List[str]], default None*) – Tags to apply to the pushed image in ECR (e.g., ["1.3.4", "latest"]). If None, defaults to ["latest"].
- **region_name** (*str, default "us-west-2"*) – AWS region containing the target ECR registry
- **ignore_docker_config** (*bool, default False*) – If True, creates a temporary Docker config to prevent using stored credentials
- **max_retries** (*int, default 3*) – Maximum number of retry attempts for failed push operations

Raises

- **ValueError** – If processing_step_id cannot be mapped to an ECR repository name, or if push operations encounter errors after all retries
- **docker.errors.APIError** – If Docker API operations fail
- **boto3.exceptions.ClientError** – If AWS ECR operations fail

Return type
None

Classes

DockerConfigManager([override_default_config])	Context manager object, suitable for use with docker-py DockerClient.login

libera_utils.aws.ecr_upload.DockerConfigManager

class libera_utils.aws.ecr_upload.**DockerConfigManager**(*override_default_config: bool = False*)

Bases: object

Context manager object, suitable for use with docker-py DockerClient.login

If override_default_config is True, dockercfg_path points to a temporary directory with a blank config. Otherwise, dockercfg_path is None, which allows DockerClient.login to use the default config location.

__init__(*override_default_config: bool = False*)

class libera_utils.aws.ecr_upload.**DockerConfigManager**(*override_default_config: bool = False*)

Context manager object, suitable for use with docker-py DockerClient.login

If override_default_config is True, dockercfg_path points to a temporary directory with a blank config. Otherwise, dockercfg_path is None, which allows DockerClient.login to use the default config location.

libera_utils.aws.ecr_upload.**_get_fresh_ecr_auth**(*region_name: str*) → dict

Get fresh ECR authentication configuration.

Parameters
region_name (*str*) – AWS region name

Returns
Authentication configuration for Docker API

Return type
dict

libera_utils.aws.ecr_upload.**_push_single_tag**(*docker_client: DockerClient*, *local_image: Image*, *full_ecr_tag: str*, *region_name: str*, *max_retries: int = 3*) → None

Push a single tagged image to ECR with retry logic and fresh authentication.

Parameters

- **docker_client** (*docker.DockerClient*) – Docker client instance
- **local_image** (*docker.models.images.Image*) – Local Docker image to push
- **full_ecr_tag** (*str*) – Complete ECR tag (registry/repository:tag)
- **region_name** (*str*) – AWS region name
- **max_retries** (*int*) – Maximum retry attempts

libera_utils.aws.ecr_upload.**build_docker_image**(*context_dir: str | Path*, *image_name: str*, *tag: str = 'latest'*, *target: str | None = None*, *platform: str = 'linux/amd64'*) → None

Build a Docker image from a specified directory and tag it with a custom name.

Parameters

- **context_dir** (*Union[str, Path]*) – The path to the directory containing the Dockerfile and other build context.
- **image_name** (*str*) – The name to give the Docker image.
- **tag** (*str, optional*) – The tag to apply to the image (default is 'latest').
- **target** (*Optional[str]*) – Name of the target to build.
- **platform** (*str*) – Default "linux/amd64".

Raises
ValueError – If the specified directory does not exist or the build fails.

libera_utils.aws.ecr_upload.**ecr_upload_cli_handler**(*parsed_args: Namespace*) → None

CLI handler function for ecr-upload CLI subcommand.

Parameters
parsed_args (*argparse.Namespace*) – Namespace of parsed CLI arguments

Return type
None

libera_utils.aws.ecr_upload.**push_image_to_ecr**(*image_name: str*, *image_tag: str*, *processing_step_id: str | ProcessingStepIdentifier*, ***, *ecr_image_tags: list[str] = None*, *region_name: str = 'us-west-2'*, *ignore_docker_config: bool = False*, *max_retries: int = 1*) → None

Push a Docker image to Amazon ECR with robust authentication handling.

This function handles ECR authentication by obtaining fresh credentials for each push operation, preventing authentication token expiration issues during multi-tag pushes.

Parameters

- **image_name** (*str*) – Local name of the Docker image
- **image_tag** (*str*) – Local tag of the Docker image (often 'latest')
- **processing_step_id** (*Union[str, ProcessingStepIdentifier]*) – Processing step ID string or object used to determine ECR repository name. L0 processing step IDs are not supported as they have no associated ECR.
- **ecr_image_tags** (*Optional[List[str]], default None*) – Tags to apply to the pushed image in ECR (e.g., ["1.3.4", "latest"]). If None, defaults to ["latest"].
- **region_name** (*str, default "us-west-2"*) – AWS region containing the target ECR registry
- **ignore_docker_config** (*bool, default False*) – If True, creates a temporary Docker config to prevent using stored credentials
- **max_retries** (*int, default 3*) – Maximum number of retry attempts for failed push operations

Raises

- **ValueError** – If processing_step_id cannot be mapped to an ECR repository name, or if push operations encounter errors after all retries
- **docker.errors.APIError** – If Docker API operations fail
- **boto3.exceptions.ClientError** – If AWS ECR operations fail

Return type
None

libera_utils.aws.processing_step_function_trigger

Module for manually triggering a step function

Functions

get_stepfunction_execution_url(execution_arn)	Generate AWS Console URL for Step Function execution
step_function_trigger(algorithm_name, ...[, ...])	Start a stepfunction to process a certain days data :param algorithm_name: The name of the algorithm to run as identified by the processing step identifier :type algorithm_name: Union[str, ProcessingStepIdentifier] :param applicable_day: The day to process data for :type applicable_day: Union[str, datetime] :param wait_time: The time to wait between checking the status of the step function, default is 5 seconds for a maximum of ~30 second total wait time :type wait_time: int, optional
step_function_trigger_cli_handler(parsed_args)	CLI handler function for the step function trigger CLI subcommand.

libera_utils.aws.processing_step_function_trigger.get_stepfunction_execution_url

libera_utils.aws.processing_step_function_trigger.**get_stepfunction_execution_url**(*execution_arn*) → str

Generate AWS Console URL for Step Function execution

libera_utils.aws.processing_step_function_trigger.step_function_trigger

libera_utils.aws.processing_step_function_trigger.**step_function_trigger**(*algorithm_name: str | ProcessingStepIdentifier*, *applicable_day: str | datetime*, *wait_time: int = 0*) → str

Start a stepfunction to process a certain days data :param algorithm_name: The name of the algorithm to run as identified by the processing step identifier :type algorithm_name: Union[str, ProcessingStepIdentifier] :param applicable_day: The day to process data for :type applicable_day: Union[str, datetime] :param wait_time: The time to wait between checking the status of the step function, default is 5 seconds for a maximum of

~30 second total wait time

Returns
The resulting status of execution. Returns N/A if n

Return type
str

libera_utils.aws.processing_step_function_trigger.step_function_trigger_cli_handler

libera_utils.aws.processing_step_function_trigger.**step_function_trigger_cli_handler**(*parsed_args: Namespace*)

CLI handler function for the step function trigger CLI subcommand.

Parameters
parsed_args (*argparse.Namespace*) – The parsed object of CLI arguments

libera_utils.aws.processing_step_function_trigger.**get_stepfunction_execution_url**(*execution_arn*) → str

Generate AWS Console URL for Step Function execution

libera_utils.aws.processing_step_function_trigger.**step_function_trigger**(*algorithm_name: str | ProcessingStepIdentifier*, *applicable_day: str | datetime*, *wait_time: int = 0*) → str

Start a stepfunction to process a certain days data :param algorithm_name: The name of the algorithm to run as identified by the processing step identifier :type algorithm_name: Union[str, ProcessingStepIdentifier] :param applicable_day: The day to process data for :type applicable_day: Union[str, datetime] :param wait_time: The time to wait between checking the status of the step function, default is 5 seconds for a maximum of

~30 second total wait time

Returns
The resulting status of execution. Returns N/A if n

Return type
str

libera_utils.aws.processing_step_function_trigger.**step_function_trigger_cli_handler**(*parsed_args: Namespace*)

CLI handler function for the step function trigger CLI subcommand.

Parameters
parsed_args (*argparse.Namespace*) – The parsed object of CLI arguments

libera_utils.aws.s3_utilities

Module for S3 cli utilities

Functions

s3_copy_cli_handler(parsed_args)	CLI handler function for s3-utils cp CLI subcommand.
s3_list_archive_files(processing_step, *[, ...])	List all files in an archive S3 bucket for a given processing step.
s3_list_cli_handler(parsed_args)	CLI handler function for s3-utils list CLI subcommand.
s3_put_cli_handler(parsed_args)	CLI handler function for s3-utils put CLI subcommand.
s3_put_in_archive_for_processing_step(...[, ...])	Upload a file to the archive S3 bucket associated with a given processing step.

libera_utils.aws.s3_utilities.s3_copy_cli_handler

libera_utils.aws.s3_utilities.**s3_copy_cli_handler**(*parsed_args: Namespace*) → None

CLI handler function for s3-utils cp CLI subcommand.

libera_utils.aws.s3_utilities.s3_list_archive_files

libera_utils.aws.s3_utilities.**s3_list_archive_files**(*processing_step: str | ProcessingStepIdentifier*, **, account_suffix: str | LiberaAccountSuffix | None = LiberaAccountSuffix.STAGE*) → list

List all files in an archive S3 bucket for a given processing step.

Parameters

- **processing_step** (*str*) – Processing step ID string. Used to infer the S3 archive bucket name.
- **account_suffix** (*Union[str, LiberaAccountSuffix], optional*) – Account suffix for the bucket name, by default LiberaAccountSuffix.STAGE
- **print_out** (*bool, optional*) – Print the list of files to the console, by default False

Returns

bucket_objects – S3Path objects for each file in the bucket

Return type

list

libera_utils.aws.s3_utilities.s3_list_cli_handler

libera_utils.aws.s3_utilities.**s3_list_cli_handler**(*parsed_args: Namespace*) → None

CLI handler function for s3-utils list CLI subcommand.

libera_utils.aws.s3_utilities.s3_put_cli_handler

libera_utils.aws.s3_utilities.**s3_put_cli_handler**(*parsed_args: Namespace*) → None

CLI handler function for s3-utils put CLI subcommand.

libera_utils.aws.s3_utilities.s3_put_in_archive_for_processing_step

libera_utils.aws.s3_utilities.**s3_put_in_archive_for_processing_step**(*path_to_file: Path | S3Path*, *processing_step: str | ProcessingStepIdentifier*, **, account_suffix: str | LiberaAccountSuffix | None = LiberaAccountSuffix.STAGE*)

Upload a file to the archive S3 bucket associated with a given processing step.

Parameters

- **path_to_file** (*Path*) – Local path to the file to upload
- **processing_step** (*str*) – processing_step : Union[str, ProcessingStepIdentifier] Processing step ID string or object. Used to infer the S3 archive bucket name.
- **account_suffix** (*Union[str, LiberaAccountSuffix], optional*) – Account suffix for the bucket name, by default LiberaAccountSuffix.STAGE

libera_utils.aws.s3_utilities.**s3_copy_cli_handler**(*parsed_args: Namespace*) → None

CLI handler function for s3-utils cp CLI subcommand.

libera_utils.aws.s3_utilities.**s3_list_archive_files**(*processing_step: str | ProcessingStepIdentifier, *, account_suffix: str | LiberaAccountSuffix | None = LiberaAccountSuffix.STAGE*) → list

List all files in an archive S3 bucket for a given processing step.

Parameters

- **processing_step** (*str*) – Processing step ID string. Used to infer the S3 archive bucket name.
- **account_suffix** (*Union[str, LiberaAccountSuffix], optional*) – Account suffix for the bucket name, by default LiberaAccountSuffix.STAGE
- **print_out** (*bool, optional*) – Print the list of files to the console, by default False

Returns

bucket_objects – S3Path objects for each file in the bucket

Return type

list

libera_utils.aws.s3_utilities.**s3_list_cli_handler**(*parsed_args: Namespace*) → None

CLI handler function for s3-utils list CLI subcommand.

libera_utils.aws.s3_utilities.**s3_put_cli_handler**(*parsed_args: Namespace*) → None

CLI handler function for s3-utils put CLI subcommand.

libera_utils.aws.s3_utilities.**s3_put_in_archive_for_processing_step**(*path_to_file: Path | S3Path, processing_step: str | ProcessingStepIdentifier, *, account_suffix: str | LiberaAccountSuffix | None = LiberaAccountSuffix.STAGE*)

Upload a file to the archive S3 bucket associated with a given processing step.

Parameters

- **path_to_file** (*Path*) – Local path to the file to upload
- **processing_step** (*str*) – processing_step : Union[str, ProcessingStepIdentifier] Processing step ID string or object. Used to infer the S3 archive bucket name.
- **account_suffix** (*Union[str, LiberaAccountSuffix], optional*) – Account suffix for the bucket name, by default LiberaAccountSuffix.STAGE

libera_utils.aws.utils

Helper functions for AWS access

Functions

get_aws_account_number([region_name])	Get a users AWS account ID number :param region_name: Region that the users AWS account is on :type region_name: string

libera_utils.aws.utils.get_aws_account_number

libera_utils.aws.utils.**get_aws_account_number**(*region_name='us-west-2'*)

Get a users AWS account ID number :param region_name: Region that the users AWS account is on :type region_name: string

Returns
account_id – users account_id number

Return type
int

libera_utils.aws.utils.**get_aws_account_number**(*region_name='us-west-2'*)

Get a users AWS account ID number :param region_name: Region that the users AWS account is on :type region_name: string

Returns
account_id – users account_id number

Return type
int

3.1.2 libera_utils.cli

Module for the Libera SDC utilities CLI

Functions

main([cli_args])	Main CLI entrypoint that runs the function inferred from the specified subcommand
parse_cli_args(cli_args)	Parse CLI arguments
print_version_info(*args)	Print CLI version information

libera_utils.cli.main

libera_utils.cli.**main**(*cli_args: list = None*)

Main CLI entrypoint that runs the function inferred from the specified subcommand

libera_utils.cli.parse_cli_args

libera_utils.cli.**parse_cli_args**(*cli_args: list*)

Parse CLI arguments

Parameters
cli_args (*list*) – List of CLI arguments to parse

Returns
Parsed arguments in a Namespace object

Return type
argparse.Namespace

libera_utils.cli.print_version_info

libera_utils.cli.**print_version_info**(**args*)

Print CLI version information

libera_utils.cli.**main**(*cli_args: list = None*)

Main CLI entrypoint that runs the function inferred from the specified subcommand

libera_utils.cli.**parse_cli_args**(*cli_args: list*)

Parse CLI arguments

Parameters
cli_args (*list*) – List of CLI arguments to parse

Returns
Parsed arguments in a Namespace object

Return type
argparse.Namespace

libera_utils.cli.**print_version_info**(**args*)

Print CLI version information

3.1.3 libera_utils.config

Configuration reader. To modify the configuration, see file: config.json

Module Attributes

config	Singleton (one per process) accessor for *libera_utils.config._ConfigurationCache()*

libera_utils.config.config

libera_utils.config.**config = <libera_utils.config._ConfigurationCache object>**

Singleton (one per process) accessor for *libera_utils.config._ConfigurationCache()*

Classes

ConfigurationFormatter()	Customize the string formatter to replace fields in a config string with values from the configuration dictionary.

libera_utils.config.ConfigurationFormatter

class libera_utils.config.**ConfigurationFormatter**

Bases: Formatter

Customize the string formatter to replace fields in a config string with values from the configuration dictionary. This will allow configuration parameters in the emus_config.json file to be based off of other configuration parameters by wrapping the configuration key in curly braces.

Methods

get_field(field_name, args, kwargs)	
get_value(key, *args, **kwargs)	Overrides the default get_value method in the python formatter.
parse(format_string)	

check_unused_args
convert_field
format
format_field
vformat

__init__(**args*, ***kwargs*)

Methods

get_value(key, *args, **kwargs)	Overrides the default get_value method in the python formatter.

get_value(*key: str*, **args*, ***kwargs*)

Overrides the default get_value method in the python formatter. This will return the value from the emus configuration with the specified key.

class libera_utils.config.**ConfigurationFormatter**

Customize the string formatter to replace fields in a config string with values from the configuration dictionary. This will allow configuration parameters in the emus_config.json file to be based off of other configuration parameters by wrapping the configuration key in curly braces.

Methods

get_field(field_name, args, kwargs)	
get_value(key, *args, **kwargs)	Overrides the default get_value method in the python formatter.
parse(format_string)	

check_unused_args
convert_field
format
format_field
vformat

get_value(*key: str*, **args*, ***kwargs*)

Overrides the default get_value method in the python formatter. This will return the value from the emus configuration with the specified key.

class libera_utils.config.**_ConfigurationCache**

Class that stores the JSON configuration and provides methods for accessing configuration information

Methods

force_reload()	Force reloading of the JSON config
get(key)	Retrieves a configuration value from either the cached JSON or from the environment

_format_return_value(*value: str*)

Recursively formats the returned value, looking for config keys to substitute.

Parameters
value (*str*) – String to format

Return type
str

_parse_numeric_types(*value: str*)

Checks the final result of a config retrieval. If it is a string that can be interpreted as a float or int, parse it and return that.

Parameters
value (*any*) – Final formatted value.

Return type
str or float or int

force_reload()

Force reloading of the JSON config

get(*key*)

Retrieves a configuration value from either the cached JSON or from the environment

Parameters
key (*str*) – Key for which to retrieve the configured value.

Returns
Resulting value

Return type
any

libera_utils.config.**config = <libera_utils.config._ConfigurationCache object>**

Singleton (one per process) accessor for *libera_utils.config._ConfigurationCache()*

3.1.4 libera_utils.constants

Constants module used throughout the libera_utils package

Classes

DataLevel(value)	Data product level
DataProductIdentifier(value)	Enumeration of data product canonical IDs used in AWS resource naming.
LiberaApid(value)	APIDs for packets
ManifestType(value)	Enumerated legal manifest type values
ProcessingStepIdentifier(value)	Enumeration of processing step IDs used in AWS resource naming and processing orchestration.

libera_utils.constants.DataLevel

class libera_utils.constants.**DataLevel**(*value*)

Bases: StrEnum

Data product level

Methods

capitalize(/)	Return a capitalized version of the string.
casefold(/)	Return a version of the string suitable for caseless comparisons.
center(width[, fillchar])	Return a centered string of length width.
count(sub[, start[, end]])	Return the number of non-overlapping occurrences of substring sub in string S[start:end].
encode(/[, encoding, errors])	Encode the string using the codec registered for encoding.
endswith(suffix[, start[, end]])	Return True if S ends with the specified suffix, False otherwise.
expandtabs(/[, tabsize])	Return a copy where all tab characters are expanded using spaces.
find(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
format(*args, **kwargs)	Return a formatted version of S, using substitutions from args and kwargs.
format_map(mapping)	Return a formatted version of S, using substitutions from mapping.
index(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
isalnum(/)	Return True if the string is an alpha-numeric string, False otherwise.
isalpha(/)	Return True if the string is an alphabetic string, False otherwise.
isascii(/)	Return True if all characters in the string are ASCII, False otherwise.
isdecimal(/)	Return True if the string is a decimal string, False otherwise.
isdigit(/)	Return True if the string is a digit string, False otherwise.
isidentifier(/)	Return True if the string is a valid Python identifier, False otherwise.

continues on next page

Table 29 – continued from previous page

islower(/)	Return True if the string is a lowercase string, False otherwise.
isnumeric(/)	Return True if the string is a numeric string, False otherwise.
isprintable(/)	Return True if the string is printable, False otherwise.
isspace(/)	Return True if the string is a whitespace string, False otherwise.
istitle(/)	Return True if the string is a title-cased string, False otherwise.
isupper(/)	Return True if the string is an uppercase string, False otherwise.
join(iterable, /)	Concatenate any number of strings.
ljust(width[, fillchar])	Return a left-justified string of length width.
lower(/)	Return a copy of the string converted to lowercase.
lstrip([chars])	Return a copy of the string with leading whitespace removed.
maketrans(x[, y, z])	Return a translation table usable for str.translate().
partition(sep, /)	Partition the string into three parts using the given separator.
removeprefix(prefix, /)	Return a str with the given prefix string removed if present.
removesuffix(suffix, /)	Return a str with the given suffix string removed if present.
replace(old, new[, count])	Return a copy with all occurrences of substring old replaced by new.
rfind(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rindex(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rjust(width[, fillchar])	Return a right-justified string of length width.
rpartition(sep, /)	Partition the string into three parts using the given separator.
rsplit(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
rstrip([chars])	Return a copy of the string with trailing whitespace removed.
split(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
splitlines(/[, keepends])	Return a list of the lines in the string, breaking at line boundaries.
startswith(prefix[, start[, end]])	Return True if S starts with the specified prefix, False otherwise.
strip([chars])	Return a copy of the string with leading and trailing whitespace removed.
swapcase(/)	Convert uppercase characters to lowercase and lowercase characters to uppercase.
title(/)	Return a version of the string where each word is titlecased.
translate(table, /)	Replace each character in the string using the given translation table.
upper(/)	Return a copy of the string converted to uppercase.

continues on next page

Table 29 – continued from previous page

zfill(width, /)	Pad a numeric string with zeros on the left, to fill a field of the given width.

__init__(**args*, ***kwds*)

Methods

encode(/[, encoding, errors])	Encode the string using the codec registered for encoding.
replace(old, new[, count])	Return a copy with all occurrences of substring old replaced by new.
split(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
rsplit(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
join(iterable, /)	Concatenate any number of strings.
capitalize(/)	Return a capitalized version of the string.
casefold(/)	Return a version of the string suitable for caseless comparisons.
title(/)	Return a version of the string where each word is titlecased.
center(width[, fillchar])	Return a centered string of length width.
count(sub[, start[, end]])	Return the number of non-overlapping occurrences of substring sub in string S[start:end].
expandtabs(/[, tabsize])	Return a copy where all tab characters are expanded using spaces.
find(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
partition(sep, /)	Partition the string into three parts using the given separator.
index(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
ljust(width[, fillchar])	Return a left-justified string of length width.
lower(/)	Return a copy of the string converted to lowercase.
lstrip([chars])	Return a copy of the string with leading whitespace removed.
rfind(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rindex(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rjust(width[, fillchar])	Return a right-justified string of length width.
rstrip([chars])	Return a copy of the string with trailing whitespace removed.
rpartition(sep, /)	Partition the string into three parts using the given separator.
splitlines(/[, keepends])	Return a list of the lines in the string, breaking at line boundaries.
strip([chars])	Return a copy of the string with leading and trailing whitespace removed.

continues on next page

Table 30 – continued from previous page

swapcase(/)	Convert uppercase characters to lowercase and lowercase characters to uppercase.
translate(table, /)	Replace each character in the string using the given translation table.
upper(/)	Return a copy of the string converted to uppercase.
startswith(prefix[, start[, end]])	Return True if S starts with the specified prefix, False otherwise.
endswith(suffix[, start[, end]])	Return True if S ends with the specified suffix, False otherwise.
removeprefix(prefix, /)	Return a str with the given prefix string removed if present.
removesuffix(suffix, /)	Return a str with the given suffix string removed if present.
isascii(/)	Return True if all characters in the string are ASCII, False otherwise.
islower(/)	Return True if the string is a lowercase string, False otherwise.
isupper(/)	Return True if the string is an uppercase string, False otherwise.
istitle(/)	Return True if the string is a title-cased string, False otherwise.
isspace(/)	Return True if the string is a whitespace string, False otherwise.
isdecimal(/)	Return True if the string is a decimal string, False otherwise.
isdigit(/)	Return True if the string is a digit string, False otherwise.
isnumeric(/)	Return True if the string is a numeric string, False otherwise.
isalpha(/)	Return True if the string is an alphabetic string, False otherwise.
isalnum(/)	Return True if the string is an alpha-numeric string, False otherwise.
isidentifier(/)	Return True if the string is a valid Python identifier, False otherwise.
isprintable(/)	Return True if the string is printable, False otherwise.
zfill(width, /)	Pad a numeric string with zeros on the left, to fill a field of the given width.
format(*args, **kwargs)	Return a formatted version of S, using substitutions from args and kwargs.
format_map(mapping)	Return a formatted version of S, using substitutions from mapping.
maketrans(x[, y, z])	Return a translation table usable for str.translate().

Attributes

archive_bucket_name	Gets the archive bucket name for the data level.
`L0`	
`SPICE`	

continues on next page

Table 31 – continued from previous page

CAL
L1A
L1B
L2
ANC

property archive_bucket_name: str

Gets the archive bucket name for the data level.

Notes

This does not include any account suffix, which must be added separately.

capitalize(/)

Return a capitalized version of the string.

More specifically, make the first character have upper case and the rest lower case.

casefold(/)

Return a version of the string suitable for caseless comparisons.

center(*width*, *fillchar=' '*, /)

Return a centered string of length width.

Padding is done using the specified fill character (default is a space).

count(*sub*[, *start*[, *end*]]) → int

Return the number of non-overlapping occurrences of substring sub in string S[start:end]. Optional arguments start and end are interpreted as in slice notation.

encode(/, *encoding='utf-8'*, *errors='strict'*)

Encode the string using the codec registered for encoding.

encoding
The encoding in which to encode the string.

errors
The error handling scheme to use for encoding errors. The default is 'strict' meaning that encoding errors raise a UnicodeEncodeError. Other possible values are 'ignore', 'replace' and 'xmlcharrefreplace' as well as any other name registered with codecs.register_error that can handle UnicodeEncodeErrors.

endswith(*suffix*[, *start*[, *end*]]) → bool

Return True if S ends with the specified suffix, False otherwise. With optional start, test S beginning at that position. With optional end, stop comparing S at that position. suffix can also be a tuple of strings to try.

expandtabs(/, *tabsize=8*)

Return a copy where all tab characters are expanded using spaces.

If tabsize is not given, a tab size of 8 characters is assumed.

find(*sub*[, *start*[, *end*]]) → int

Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end]. Optional arguments start and end are interpreted as in slice notation.

Return -1 on failure.

format(**args*, ***kwargs*) → str

Return a formatted version of S, using substitutions from args and kwargs. The substitutions are identified by braces ('{' and '}').

format_map(*mapping*) → str

Return a formatted version of S, using substitutions from mapping. The substitutions are identified by braces ('{' and '}').

index(*sub*[, *start*[, *end*]]) → int

Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end]. Optional arguments start and end are interpreted as in slice notation.

Raises ValueError when the substring is not found.

isalnum(/)

Return True if the string is an alpha-numeric string, False otherwise.

A string is alpha-numeric if all characters in the string are alpha-numeric and there is at least one character in the string.

isalpha(/)

Return True if the string is an alphabetic string, False otherwise.

A string is alphabetic if all characters in the string are alphabetic and there is at least one character in the string.

isascii(/)

Return True if all characters in the string are ASCII, False otherwise.

ASCII characters have code points in the range U+0000-U+007F. Empty string is ASCII too.

isdecimal(/)

Return True if the string is a decimal string, False otherwise.

A string is a decimal string if all characters in the string are decimal and there is at least one character in the string.

isdigit(/)

Return True if the string is a digit string, False otherwise.

A string is a digit string if all characters in the string are digits and there is at least one character in the string.

isidentifier(/)

Return True if the string is a valid Python identifier, False otherwise.

Call keyword.iskeyword(s) to test whether string s is a reserved identifier, such as "def" or "class".

islower(/)

Return True if the string is a lowercase string, False otherwise.

A string is lowercase if all cased characters in the string are lowercase and there is at least one cased character in the string.

isnumeric(/)

Return True if the string is a numeric string, False otherwise.

A string is numeric if all characters in the string are numeric and there is at least one character in the string.

isprintable(/)

Return True if the string is printable, False otherwise.

A string is printable if all of its characters are considered printable in repr() or if it is empty.

isspace(/)

Return True if the string is a whitespace string, False otherwise.

A string is whitespace if all characters in the string are whitespace and there is at least one character in the string.

istitle(/)

Return True if the string is a title-cased string, False otherwise.

In a title-cased string, upper- and title-case characters may only follow uncased characters and lowercase characters only cased ones.

isupper(/)

Return True if the string is an uppercase string, False otherwise.

A string is uppercase if all cased characters in the string are uppercase and there is at least one cased character in the string.

join(*iterable*, /)

Concatenate any number of strings.

The string whose method is called is inserted in between each given string. The result is returned as a new string.

Example: '.'.join(['ab', 'pq', 'rs']) -> 'ab.pq.rs'

ljust(*width*, *fillchar=' '*, /)

Return a left-justified string of length width.

Padding is done using the specified fill character (default is a space).

lower(/)

Return a copy of the string converted to lowercase.

lstrip(*chars=None*, /)

Return a copy of the string with leading whitespace removed.

If chars is given and not None, remove characters in chars instead.

static maketrans(*x*, *y=<unrepresentable>*, *z=<unrepresentable>*, /)

Return a translation table usable for str.translate().

If there is only one argument, it must be a dictionary mapping Unicode ordinals (integers) or characters to Unicode ordinals, strings or None. Character keys will be then converted to ordinals. If there are two arguments, they must be strings of equal length, and in the resulting dictionary, each character in x will be mapped to the character at the same position in y. If there is a third argument, it must be a string, whose characters will be mapped to None in the result.

partition(*sep*, /)

Partition the string into three parts using the given separator.

This will search for the separator in the string. If the separator is found, returns a 3-tuple containing the part before the separator, the separator itself, and the part after it.

If the separator is not found, returns a 3-tuple containing the original string and two empty strings.

removeprefix(*prefix*, /)

Return a str with the given prefix string removed if present.

If the string starts with the prefix string, return string[len(prefix):]. Otherwise, return a copy of the original string.

removesuffix(*suffix*, /)

Return a str with the given suffix string removed if present.

If the string ends with the suffix string and that suffix is not empty, return string[:-len(suffix)]. Otherwise, return a copy of the original string.

replace(*old*, *new*, *count=-1*, /)

Return a copy with all occurrences of substring old replaced by new.

count
Maximum number of occurrences to replace. -1 (the default value) means replace all occurrences.

If the optional argument count is given, only the first count occurrences are replaced.

rfind(*sub*[, *start*[, *end*]]) → int

Return the highest index in S where substring sub is found, such that sub is contained within S[start:end]. Optional arguments start and end are interpreted as in slice notation.

Return -1 on failure.

rindex(*sub*[, *start*[, *end*]]) → int

Return the highest index in S where substring sub is found, such that sub is contained within S[start:end]. Optional arguments start and end are interpreted as in slice notation.

Raises ValueError when the substring is not found.

rjust(*width*, *fillchar=' '*, /)

Return a right-justified string of length width.

Padding is done using the specified fill character (default is a space).

rpartition(*sep*, /)

Partition the string into three parts using the given separator.

This will search for the separator in the string, starting at the end. If the separator is found, returns a 3-tuple containing the part before the separator, the separator itself, and the part after it.

If the separator is not found, returns a 3-tuple containing two empty strings and the original string.

rsplit(/, *sep=None*, *maxsplit=-1*)

Return a list of the substrings in the string, using sep as the separator string.

sep
The separator used to split the string.

When set to None (the default value), will split on any whitespace character (including n r t f and spaces) and will discard empty strings from the result.

maxsplit
Maximum number of splits. -1 (the default value) means no limit.

Splitting starts at the end of the string and works to the front.

`rstrip`(*chars=None*, */*)

Return a copy of the string with trailing whitespace removed.

If chars is given and not None, remove characters in chars instead.

`split`(*/*, *sep=None*, *maxsplit=-1*)

Return a list of the substrings in the string, using sep as the separator string.

sep
The separator used to split the string.

When set to None (the default value), will split on any whitespace character (including n r t f and spaces) and will discard empty strings from the result.

maxsplit
Maximum number of splits. -1 (the default value) means no limit.

Splitting starts at the front of the string and works to the end.

Note, str.split() is mainly useful for data that has been intentionally delimited. With natural text that includes punctuation, consider using the regular expression module.

`splitlines`(*/*, *keepends=False*)

Return a list of the lines in the string, breaking at line boundaries.

Line breaks are not included in the resulting list unless keepends is given and true.

`startswith`(*prefix*[, *start*[, *end*]]) → bool

Return True if S starts with the specified prefix, False otherwise. With optional start, test S beginning at that position. With optional end, stop comparing S at that position. prefix can also be a tuple of strings to try.

`strip`(*chars=None*, */*)

Return a copy of the string with leading and trailing whitespace removed.

If chars is given and not None, remove characters in chars instead.

`swapcase`(*/*)

Convert uppercase characters to lowercase and lowercase characters to uppercase.

`title`(*/*)

Return a version of the string where each word is titlecased.

More specifically, words start with uppercased characters and all remaining cased characters have lower case.

`translate`(*table*, */*)

Replace each character in the string using the given translation table.

table
Translation table, which must be a mapping of Unicode ordinals to Unicode ordinals, strings, or None.

The table must implement lookup/indexing via __getitem__, for instance a dictionary or list. If this operation raises LookupError, the character is left untouched. Characters mapped to None are deleted.

upper(/)

Return a copy of the string converted to uppercase.

zfill(*width*, /)

Pad a numeric string with zeros on the left, to fill a field of the given width.

The string is never truncated.

libera_utils.constants.DataProductIdentifier

class libera_utils.constants.**DataProductIdentifier**(*value*)

Bases: StrEnum

Enumeration of data product canonical IDs used in AWS resource naming.

These IDs refer to the data products (files) themselves, NOT the processing steps (since processing steps may produce multiple products). The string values are the product names used in filenames.

This enum replaces the old ProductName enum from filenaming.py to provide a single source of truth.

Each member is defined as a tuple: (product_name, data_level) - product_name: The string value used in filenames and AWS resources - data_level: The DataLevel enum value indicating the processing level

Example

```
>>> product = DataProductIdentifier.l1b_rad
>>> str(product)  # Returns "RAD-4CH"
>>> product.level  # Returns DataLevel.L1B
>>> product.level.archive_bucket_name  # Returns "libera-l1b-data"
```

When adding new products:

1. Add the enum member with its product name and DataLevel
2. No need to update any lookup dictionaries - metadata is embedded!

Methods

capitalize(/)	Return a capitalized version of the string.
casefold(/)	Return a version of the string suitable for caseless comparisons.
center(width[, fillchar])	Return a centered string of length width.
count(sub[, start[, end]])	Return the number of non-overlapping occurrences of substring sub in string S[start:end].
encode(/[, encoding, errors])	Encode the string using the codec registered for encoding.
endswith(suffix[, start[, end]])	Return True if S ends with the specified suffix, False otherwise.
expandtabs(/[, tabsize])	Return a copy where all tab characters are expanded using spaces.
find(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
format(*args, **kwargs)	Return a formatted version of S, using substitutions from args and kwargs.

continues on next page

Table 32 – continued from previous page

format_map(mapping)	Return a formatted version of S, using substitutions from mapping.
index(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
isalnum(/)	Return True if the string is an alpha-numeric string, False otherwise.
isalpha(/)	Return True if the string is an alphabetic string, False otherwise.
isascii(/)	Return True if all characters in the string are ASCII, False otherwise.
isdecimal(/)	Return True if the string is a decimal string, False otherwise.
isdigit(/)	Return True if the string is a digit string, False otherwise.
isidentifier(/)	Return True if the string is a valid Python identifier, False otherwise.
islower(/)	Return True if the string is a lowercase string, False otherwise.
isnumeric(/)	Return True if the string is a numeric string, False otherwise.
isprintable(/)	Return True if the string is printable, False otherwise.
isspace(/)	Return True if the string is a whitespace string, False otherwise.
istitle(/)	Return True if the string is a title-cased string, False otherwise.
isupper(/)	Return True if the string is an uppercase string, False otherwise.
join(iterable, /)	Concatenate any number of strings.
ljust(width[, fillchar])	Return a left-justified string of length width.
lower(/)	Return a copy of the string converted to lowercase.
lstrip([chars])	Return a copy of the string with leading whitespace removed.
maketrans(x[, y, z])	Return a translation table usable for str.translate().
partition(sep, /)	Partition the string into three parts using the given separator.
removeprefix(prefix, /)	Return a str with the given prefix string removed if present.
removesuffix(suffix, /)	Return a str with the given suffix string removed if present.
replace(old, new[, count])	Return a copy with all occurrences of substring old replaced by new.
rfind(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rindex(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rjust(width[, fillchar])	Return a right-justified string of length width.
rpartition(sep, /)	Partition the string into three parts using the given separator.
rsplit(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
rstrip([chars])	Return a copy of the string with trailing whitespace removed.

continues on next page

Table 32 – continued from previous page

split(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
splitlines(/[, keepends])	Return a list of the lines in the string, breaking at line boundaries.
startswith(prefix[, start[, end]])	Return True if S starts with the specified prefix, False otherwise.
strip([chars])	Return a copy of the string with leading and trailing whitespace removed.
swapcase(/)	Convert uppercase characters to lowercase and lowercase characters to uppercase.
title(/)	Return a version of the string where each word is titlecased.
translate(table, /)	Replace each character in the string using the given translation table.
upper(/)	Return a copy of the string converted to uppercase.
zfill(width, /)	Pad a numeric string with zeros on the left, to fill a field of the given width.

__init__(**args*, ***kwds*)

Methods

encode(/[, encoding, errors])	Encode the string using the codec registered for encoding.
replace(old, new[, count])	Return a copy with all occurrences of substring old replaced by new.
split(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
rsplit(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
join(iterable, /)	Concatenate any number of strings.
capitalize(/)	Return a capitalized version of the string.
casefold(/)	Return a version of the string suitable for caseless comparisons.
title(/)	Return a version of the string where each word is titlecased.
center(width[, fillchar])	Return a centered string of length width.
count(sub[, start[, end]])	Return the number of non-overlapping occurrences of substring sub in string S[start:end].
expandtabs(/[, tabsize])	Return a copy where all tab characters are expanded using spaces.
find(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
partition(sep, /)	Partition the string into three parts using the given separator.
index(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
ljust(width[, fillchar])	Return a left-justified string of length width.
lower(/)	Return a copy of the string converted to lowercase.

continues on next page

Table 33 – continued from previous page

lstrip([chars])	Return a copy of the string with leading whitespace removed.
rfind(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rindex(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rjust(width[, fillchar])	Return a right-justified string of length width.
rstrip([chars])	Return a copy of the string with trailing whitespace removed.
rpartition(sep, /)	Partition the string into three parts using the given separator.
splitlines(/[, keepends])	Return a list of the lines in the string, breaking at line boundaries.
strip([chars])	Return a copy of the string with leading and trailing whitespace removed.
swapcase(/)	Convert uppercase characters to lowercase and lowercase characters to uppercase.
translate(table, /)	Replace each character in the string using the given translation table.
upper(/)	Return a copy of the string converted to uppercase.
startswith(prefix[, start[, end]])	Return True if S starts with the specified prefix, False otherwise.
endswith(suffix[, start[, end]])	Return True if S ends with the specified suffix, False otherwise.
removeprefix(prefix, /)	Return a str with the given prefix string removed if present.
removesuffix(suffix, /)	Return a str with the given suffix string removed if present.
isascii(/)	Return True if all characters in the string are ASCII, False otherwise.
islower(/)	Return True if the string is a lowercase string, False otherwise.
isupper(/)	Return True if the string is an uppercase string, False otherwise.
istitle(/)	Return True if the string is a title-cased string, False otherwise.
isspace(/)	Return True if the string is a whitespace string, False otherwise.
isdecimal(/)	Return True if the string is a decimal string, False otherwise.
isdigit(/)	Return True if the string is a digit string, False otherwise.
isnumeric(/)	Return True if the string is a numeric string, False otherwise.
isalpha(/)	Return True if the string is an alphabetic string, False otherwise.
isalnum(/)	Return True if the string is an alpha-numeric string, False otherwise.
isidentifier(/)	Return True if the string is a valid Python identifier, False otherwise.
isprintable(/)	Return True if the string is printable, False otherwise.

continues on next page

Table 33 – continued from previous page

zfill(width, /)	Pad a numeric string with zeros on the left, to fill a field of the given width.
format(*args, **kwargs)	Return a formatted version of S, using substitutions from args and kwargs.
format_map(mapping)	Return a formatted version of S, using substitutions from mapping.
maketrans(x[, y, z])	Return a translation table usable for str.translate().
get_partial_archive_bucket_name()	Gets the archive bucket name from the data product identifier .

Attributes

product_name	Get the name formatted for AWS resources for this data product
data_level	Get the processing level for this data product.
l0_pds_cr	
l0_jpss_sc_pos_pds	
l0_icie_rad_sample_pds	
l0_icie_wfov_sci_pds	
l0_icie_axis_sample_pds	
l0_icie_sw_stat_pds	
l0_icie_seq_hk_pds	
l0_icie_fp_hk_pds	
l0_icie_log_msg_pds	
l0_icie_rad_full_pds	
l0_icie_axis_hk_pds	
l0_icie_wfov_hk_pds	
l0_icie_cal_full_pds	
l0_icie_cal_sample_pds	
l0_icie_wfov_resp_pds	
l0_icie_crit_hk_pds	
l0_icie_nom_hk_pds	
l0_icie_ana_hk_pds	

continues on next page

Table 34 – continued from previous page

l0_icie_temp_hk_pds
l1a_jpss_sc_pos_decoded
l1a_icie_rad_sample_decoded
l1a_icie_wfov_sci_decoded
l1a_icie_axis_sample_decoded
l1a_icie_sw_stat_decoded
l1a_icie_seq_hk_decoded
l1a_icie_fp_hk_decoded
l1a_icie_log_msg_decoded
l1a_icie_rad_full_decoded
l1a_icie_axis_hk_decoded
l1a_icie_wfov_hk_decoded
l1a_icie_cal_full_decoded
l1a_icie_cal_sample_decoded
l1a_icie_wfov_resp_decoded
l1a_icie_crit_hk_decoded
l1a_icie_nom_hk_decoded
l1a_icie_ana_hk_decoded
l1a_icie_temp_hk_decoded
spice_az_ck
spice_el_ck
spice_jpss_ck
spice_jpss_spk
cal_rad
cal_cam
l1b_rad

continues on next page

Table 34 – continued from previous page

l1b_cam
l2_unf
l2_cf_rad
l2_cf_cam
l2_ssw_toa_osse
l2_ssw_toa_erbe
l2_ssw_toa_trmm
l2_ssw_toa_rt
l2_ssw_surf
anc_adm
anc_scene_id

capitalize(/)

Return a capitalized version of the string.

More specifically, make the first character have upper case and the rest lower case.

casefold(/)

Return a version of the string suitable for caseless comparisons.

center(*width*, *fillchar=' '*, /)

Return a centered string of length width.

Padding is done using the specified fill character (default is a space).

count(*sub*[, *start*[, *end*]]) → int

Return the number of non-overlapping occurrences of substring sub in string S[start:end]. Optional arguments start and end are interpreted as in slice notation.

property data_level: *DataLevel*

Get the processing level for this data product.

Returns
The processing level of this data product

Return type
DataLevel

encode(/, *encoding='utf-8'*, *errors='strict'*)

Encode the string using the codec registered for encoding.

encoding
The encoding in which to encode the string.

errors
The error handling scheme to use for encoding errors. The default is 'strict' meaning that encoding errors raise a UnicodeEncodeError. Other possible values are 'ignore', 'replace' and 'xmlcharrefreplace' as well as any other name registered with codecs.register_error that can handle UnicodeEncodeErrors.

endswith(*suffix*[, *start*[, *end*]]) → bool

Return True if S ends with the specified suffix, False otherwise. With optional start, test S beginning at that position. With optional end, stop comparing S at that position. suffix can also be a tuple of strings to try.

expandtabs(*/*, *tabsize=8*)

Return a copy where all tab characters are expanded using spaces.

If tabsize is not given, a tab size of 8 characters is assumed.

find(*sub*[, *start*[, *end*]]) → int

Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end]. Optional arguments start and end are interpreted as in slice notation.

Return -1 on failure.

format(**args*, ***kwargs*) → str

Return a formatted version of S, using substitutions from args and kwargs. The substitutions are identified by braces ('{' and '}').

format_map(*mapping*) → str

Return a formatted version of S, using substitutions from mapping. The substitutions are identified by braces ('{' and '}').

get_partial_archive_bucket_name() → str

Gets the archive bucket name from the data product identifier .

Buckets are named according to the level of data they are storing and which account they are in. This is expected to be used by the L2 developers who will most commonly be working with the stage account.

Returns
The name of the archive bucket for this data product without an account suffix

Return type
str

index(*sub*[, *start*[, *end*]]) → int

Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end]. Optional arguments start and end are interpreted as in slice notation.

Raises ValueError when the substring is not found.

isalnum(*/*)

Return True if the string is an alpha-numeric string, False otherwise.

A string is alpha-numeric if all characters in the string are alpha-numeric and there is at least one character in the string.

isalpha(*/*)

Return True if the string is an alphabetic string, False otherwise.

A string is alphabetic if all characters in the string are alphabetic and there is at least one character in the string.

isascii(/)

Return True if all characters in the string are ASCII, False otherwise.

ASCII characters have code points in the range U+0000-U+007F. Empty string is ASCII too.

isdecimal(/)

Return True if the string is a decimal string, False otherwise.

A string is a decimal string if all characters in the string are decimal and there is at least one character in the string.

isdigit(/)

Return True if the string is a digit string, False otherwise.

A string is a digit string if all characters in the string are digits and there is at least one character in the string.

isidentifier(/)

Return True if the string is a valid Python identifier, False otherwise.

Call keyword.iskeyword(s) to test whether string s is a reserved identifier, such as “def” or “class”.

islower(/)

Return True if the string is a lowercase string, False otherwise.

A string is lowercase if all cased characters in the string are lowercase and there is at least one cased character in the string.

isnumeric(/)

Return True if the string is a numeric string, False otherwise.

A string is numeric if all characters in the string are numeric and there is at least one character in the string.

isprintable(/)

Return True if the string is printable, False otherwise.

A string is printable if all of its characters are considered printable in repr() or if it is empty.

isspace(/)

Return True if the string is a whitespace string, False otherwise.

A string is whitespace if all characters in the string are whitespace and there is at least one character in the string.

istitle(/)

Return True if the string is a title-cased string, False otherwise.

In a title-cased string, upper- and title-case characters may only follow uncased characters and lowercase characters only cased ones.

isupper(/)

Return True if the string is an uppercase string, False otherwise.

A string is uppercase if all cased characters in the string are uppercase and there is at least one cased character in the string.

join(*iterable*, /)

Concatenate any number of strings.

The string whose method is called is inserted in between each given string. The result is returned as a new string.

Example: ‘.’.join([‘ab’, ‘pq’, ‘rs’]) -> ‘ab.pq.rs’

ljust(*width*, *fillchar=' '*, */*)

Return a left-justified string of length width.

Padding is done using the specified fill character (default is a space).

lower(*/*)

Return a copy of the string converted to lowercase.

lstrip(*chars=None*, */*)

Return a copy of the string with leading whitespace removed.

If chars is given and not None, remove characters in chars instead.

static maketrans(*x*, *y=<unrepresentable>*, *z=<unrepresentable>*, */*)

Return a translation table usable for str.translate().

If there is only one argument, it must be a dictionary mapping Unicode ordinals (integers) or characters to Unicode ordinals, strings or None. Character keys will be then converted to ordinals. If there are two arguments, they must be strings of equal length, and in the resulting dictionary, each character in x will be mapped to the character at the same position in y. If there is a third argument, it must be a string, whose characters will be mapped to None in the result.

partition(*sep*, */*)

Partition the string into three parts using the given separator.

This will search for the separator in the string. If the separator is found, returns a 3-tuple containing the part before the separator, the separator itself, and the part after it.

If the separator is not found, returns a 3-tuple containing the original string and two empty strings.

property product_name: str

Get the name formatted for AWS resources for this data product

The name is used to create AWS resources that are specific to the data product. This is an alias to the string value for compatibility.

removeprefix(*prefix*, */*)

Return a str with the given prefix string removed if present.

If the string starts with the prefix string, return string[len(prefix):]. Otherwise, return a copy of the original string.

removesuffix(*suffix*, */*)

Return a str with the given suffix string removed if present.

If the string ends with the suffix string and that suffix is not empty, return string[:-len(suffix)]. Otherwise, return a copy of the original string.

replace(*old*, *new*, *count=-1*, */*)

Return a copy with all occurrences of substring old replaced by new.

count
Maximum number of occurrences to replace. -1 (the default value) means replace all occurrences.

If the optional argument count is given, only the first count occurrences are replaced.

rfind(*sub*[, *start*[, *end*]]) → int

Return the highest index in S where substring sub is found, such that sub is contained within S[start:end]. Optional arguments start and end are interpreted as in slice notation.

Return -1 on failure.

rindex(*sub*[, *start*[, *end*]]) → int

Return the highest index in S where substring sub is found, such that sub is contained within S[start:end]. Optional arguments start and end are interpreted as in slice notation.

Raises ValueError when the substring is not found.

rjust(*width*, *fillchar=' '*, */*)

Return a right-justified string of length width.

Padding is done using the specified fill character (default is a space).

rpartition(*sep*, */*)

Partition the string into three parts using the given separator.

This will search for the separator in the string, starting at the end. If the separator is found, returns a 3-tuple containing the part before the separator, the separator itself, and the part after it.

If the separator is not found, returns a 3-tuple containing two empty strings and the original string.

rsplit(*/*, *sep=None*, *maxsplit=-1*)

Return a list of the substrings in the string, using sep as the separator string.

sep
The separator used to split the string.

When set to None (the default value), will split on any whitespace character (including n r t f and spaces) and will discard empty strings from the result.

maxsplit
Maximum number of splits. -1 (the default value) means no limit.

Splitting starts at the end of the string and works to the front.

rstrip(*chars=None*, */*)

Return a copy of the string with trailing whitespace removed.

If chars is given and not None, remove characters in chars instead.

split(*/*, *sep=None*, *maxsplit=-1*)

Return a list of the substrings in the string, using sep as the separator string.

sep
The separator used to split the string.

When set to None (the default value), will split on any whitespace character (including n r t f and spaces) and will discard empty strings from the result.

maxsplit
Maximum number of splits. -1 (the default value) means no limit.

Splitting starts at the front of the string and works to the end.

Note, str.split() is mainly useful for data that has been intentionally delimited. With natural text that includes punctuation, consider using the regular expression module.

splitlines(*/*, *keepends=False*)

Return a list of the lines in the string, breaking at line boundaries.

Line breaks are not included in the resulting list unless keepends is given and true.

startswith(*prefix*[, *start*[, *end*]]) → bool

Return True if S starts with the specified prefix, False otherwise. With optional start, test S beginning at that position. With optional end, stop comparing S at that position. prefix can also be a tuple of strings to try.

strip(*chars=None*, /)

Return a copy of the string with leading and trailing whitespace removed.

If chars is given and not None, remove characters in chars instead.

swapcase(/)

Convert uppercase characters to lowercase and lowercase characters to uppercase.

title(/)

Return a version of the string where each word is titlecased.

More specifically, words start with uppercased characters and all remaining cased characters have lower case.

translate(*table*, /)

Replace each character in the string using the given translation table.

table
Translation table, which must be a mapping of Unicode ordinals to Unicode ordinals, strings, or None.

The table must implement lookup/indexing via __getitem__, for instance a dictionary or list. If this operation raises LookupError, the character is left untouched. Characters mapped to None are deleted.

upper(/)

Return a copy of the string converted to uppercase.

zfill(*width*, /)

Pad a numeric string with zeros on the left, to fill a field of the given width.

The string is never truncated.

libera_utils.constants.LiberaApid

class libera_utils.constants.**LiberaApid**(*value*)

Bases: IntEnum

APIDs for packets

The enum names here should be of the form <packet-source>_<system>_<contents>. e.g. for icie_rad_sample: packet_source is Libera "ICIE", the system is the "Radiometers", and the contents is radiometer "Samples".

Notes

This is useful for identifying the data product type from the APID in an L0 filename.

The enum names (e.g. icie_seq_hk) for Libera packets here should be precisely the packet names used in FSW documents and packet definitions.

Attributes

denominator
the denominator of a rational number in lowest terms

imag
the imaginary part of a complex number

numerator
: the numerator of a rational number in lowest terms

real
: the real part of a complex number

Methods

as_integer_ratio(/)	Return integer ratio.
bit_count(/)	Number of ones in the binary representation of the absolute value of self.
bit_length(/)	Number of bits necessary to represent self in binary.
conjugate	Returns self, the complex conjugate of any int.
from_bytes(/, bytes[, byteorder, signed])	Return the integer represented by the given array of bytes.
to_bytes(/[, length, byteorder, signed])	Return an array of bytes representing an integer.

__init__(**args*, ***kwds*)

Methods

conjugate	Returns self, the complex conjugate of any int.
bit_length(/)	Number of bits necessary to represent self in binary.
bit_count(/)	Number of ones in the binary representation of the absolute value of self.
to_bytes(/[, length, byteorder, signed])	Return an array of bytes representing an integer.
from_bytes(/, bytes[, byteorder, signed])	Return the integer represented by the given array of bytes.
as_integer_ratio(/)	Return integer ratio.

Attributes

real	the real part of a complex number
imag	the imaginary part of a complex number
numerator	the numerator of a rational number in lowest terms
denominator	the denominator of a rational number in lowest terms
data_product_id	Get the DataProductIdentifier for L0 PDS files with this APID
jpss_sc_pos	
icie_sw_stat	
icie_seq_hk	
icie_fp_hk	
icie_log_msg	
icie_rad_full	

continues on next page

Table 37 – continued from previous page

icie_rad_sample
icie_axis_hk
icie_wfov_hk
icie_wfov_sci
icie_cal_full
icie_cal_sample
icie_axis_sample
icie_wfov_resp
icie_crit_hk
icie_nom_hk
icie_ana_hk
icie_temp_hk

as_integer_ratio(/)

Return integer ratio.

Return a pair of integers, whose ratio is exactly equal to the original int and with a positive denominator.

```
>>> (10).as_integer_ratio()
(10, 1)
>>> (-10).as_integer_ratio()
(-10, 1)
>>> (0).as_integer_ratio()
(0, 1)
```

bit_count(/)

Number of ones in the binary representation of the absolute value of self.

Also known as the population count.

```
>>> bin(13)
'0b1101'
>>> (13).bit_count()
3
```

bit_length(/)

Number of bits necessary to represent self in binary.

```
>>> bin(37)
'0b100101'
```

(continues on next page)

(continued from previous page)

```
>>> (37).bit_length()
6
```

conjugate()

Returns self, the complex conjugate of any int.

property data_product_id: *DataProductIdentifier*

Get the DataProductIdentifier for L0 PDS files with this APID

denominator

the denominator of a rational number in lowest terms

classmethod from_bytes(*/*, *bytes*, *byteorder='big'*, ***, *signed=False*)

Return the integer represented by the given array of bytes.

bytes
: Holds the array of bytes to convert. The argument must either support the buffer protocol or be an iterable object producing bytes. Bytes and bytearray are examples of built-in objects that support the buffer protocol.

byteorder
: The byte order used to represent the integer. If byteorder is 'big', the most significant byte is at the beginning of the byte array. If byteorder is 'little', the most significant byte is at the end of the byte array. To request the native byte order of the host system, use `sys.byteorder' as the byte order value. Default is to use 'big'.

signed
: Indicates whether two's complement is used to represent the integer.

imag

the imaginary part of a complex number

numerator

the numerator of a rational number in lowest terms

real

the real part of a complex number

to_bytes(*/*, *length=1*, *byteorder='big'*, ***, *signed=False*)

Return an array of bytes representing an integer.

length
: Length of bytes object to use. An OverflowError is raised if the integer is not representable with the given number of bytes. Default is length 1.

byteorder
: The byte order used to represent the integer. If byteorder is 'big', the most significant byte is at the beginning of the byte array. If byteorder is 'little', the most significant byte is at the end of the byte array. To request the native byte order of the host system, use `sys.byteorder' as the byte order value. Default is to use 'big'.

signed
: Determines whether two's complement is used to represent the integer. If signed is False and a negative integer is given, an OverflowError is raised.

libera_utils.constants.ManifestType

class libera_utils.constants.**ManifestType**(*value*)

Bases: StrEnum

Enumerated legal manifest type values

Methods

capitalize(/)	Return a capitalized version of the string.
casefold(/)	Return a version of the string suitable for caseless comparisons.
center(width[, fillchar])	Return a centered string of length width.
count(sub[, start[, end]])	Return the number of non-overlapping occurrences of substring sub in string S[start:end].
encode(/[, encoding, errors])	Encode the string using the codec registered for encoding.
endswith(suffix[, start[, end]])	Return True if S ends with the specified suffix, False otherwise.
expandtabs(/[, tabsize])	Return a copy where all tab characters are expanded using spaces.
find(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
format(*args, **kwargs)	Return a formatted version of S, using substitutions from args and kwargs.
format_map(mapping)	Return a formatted version of S, using substitutions from mapping.
index(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
isalnum(/)	Return True if the string is an alpha-numeric string, False otherwise.
isalpha(/)	Return True if the string is an alphabetic string, False otherwise.
isascii(/)	Return True if all characters in the string are ASCII, False otherwise.
isdecimal(/)	Return True if the string is a decimal string, False otherwise.
isdigit(/)	Return True if the string is a digit string, False otherwise.
isidentifier(/)	Return True if the string is a valid Python identifier, False otherwise.
islower(/)	Return True if the string is a lowercase string, False otherwise.
isnumeric(/)	Return True if the string is a numeric string, False otherwise.
isprintable(/)	Return True if the string is printable, False otherwise.
isspace(/)	Return True if the string is a whitespace string, False otherwise.
istitle(/)	Return True if the string is a title-cased string, False otherwise.
isupper(/)	Return True if the string is an uppercase string, False otherwise.
join(iterable, /)	Concatenate any number of strings.

continues on next page

Table 38 – continued from previous page

ljust(width[, fillchar])	Return a left-justified string of length width.
lower(/)	Return a copy of the string converted to lowercase.
lstrip([chars])	Return a copy of the string with leading whitespace removed.
maketrans(x[, y, z])	Return a translation table usable for str.translate().
partition(sep, /)	Partition the string into three parts using the given separator.
removeprefix(prefix, /)	Return a str with the given prefix string removed if present.
removesuffix(suffix, /)	Return a str with the given suffix string removed if present.
replace(old, new[, count])	Return a copy with all occurrences of substring old replaced by new.
rfind(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rindex(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rjust(width[, fillchar])	Return a right-justified string of length width.
rpartition(sep, /)	Partition the string into three parts using the given separator.
rsplit(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
rstrip([chars])	Return a copy of the string with trailing whitespace removed.
split(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
splitlines(/[, keepends])	Return a list of the lines in the string, breaking at line boundaries.
startswith(prefix[, start[, end]])	Return True if S starts with the specified prefix, False otherwise.
strip([chars])	Return a copy of the string with leading and trailing whitespace removed.
swapcase(/)	Convert uppercase characters to lowercase and lowercase characters to uppercase.
title(/)	Return a version of the string where each word is titlecased.
translate(table, /)	Replace each character in the string using the given translation table.
upper(/)	Return a copy of the string converted to uppercase.
zfill(width, /)	Pad a numeric string with zeros on the left, to fill a field of the given width.

__init__(**args*, ***kwds*)

Methods

encode(/[, encoding, errors])	Encode the string using the codec registered for encoding.
replace(old, new[, count])	Return a copy with all occurrences of substring old replaced by new.

continues on next page

Table 39 – continued from previous page

split(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
rsplit(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
join(iterable, /)	Concatenate any number of strings.
capitalize(/)	Return a capitalized version of the string.
casefold(/)	Return a version of the string suitable for caseless comparisons.
title(/)	Return a version of the string where each word is titlecased.
center(width[, fillchar])	Return a centered string of length width.
count(sub[, start[, end]])	Return the number of non-overlapping occurrences of substring sub in string S[start:end].
expandtabs(/[, tabsize])	Return a copy where all tab characters are expanded using spaces.
find(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
partition(sep, /)	Partition the string into three parts using the given separator.
index(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
ljust(width[, fillchar])	Return a left-justified string of length width.
lower(/)	Return a copy of the string converted to lowercase.
lstrip([chars])	Return a copy of the string with leading whitespace removed.
rfind(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rindex(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rjust(width[, fillchar])	Return a right-justified string of length width.
rstrip([chars])	Return a copy of the string with trailing whitespace removed.
rpartition(sep, /)	Partition the string into three parts using the given separator.
splitlines(/[, keepends])	Return a list of the lines in the string, breaking at line boundaries.
strip([chars])	Return a copy of the string with leading and trailing whitespace removed.
swapcase(/)	Convert uppercase characters to lowercase and lowercase characters to uppercase.
translate(table, /)	Replace each character in the string using the given translation table.
upper(/)	Return a copy of the string converted to uppercase.
startswith(prefix[, start[, end]])	Return True if S starts with the specified prefix, False otherwise.
endswith(suffix[, start[, end]])	Return True if S ends with the specified suffix, False otherwise.
removeprefix(prefix, /)	Return a str with the given prefix string removed if present.
removesuffix(suffix, /)	Return a str with the given suffix string removed if present.

continues on next page

Table 39 – continued from previous page

isascii(/)	Return True if all characters in the string are ASCII, False otherwise.
islower(/)	Return True if the string is a lowercase string, False otherwise.
isupper(/)	Return True if the string is an uppercase string, False otherwise.
istitle(/)	Return True if the string is a title-cased string, False otherwise.
isspace(/)	Return True if the string is a whitespace string, False otherwise.
isdecimal(/)	Return True if the string is a decimal string, False otherwise.
isdigit(/)	Return True if the string is a digit string, False otherwise.
isnumeric(/)	Return True if the string is a numeric string, False otherwise.
isalpha(/)	Return True if the string is an alphabetic string, False otherwise.
isalnum(/)	Return True if the string is an alpha-numeric string, False otherwise.
isidentifier(/)	Return True if the string is a valid Python identifier, False otherwise.
isprintable(/)	Return True if the string is printable, False otherwise.
zfill(width, /)	Pad a numeric string with zeros on the left, to fill a field of the given width.
format(*args, **kwargs)	Return a formatted version of S, using substitutions from args and kwargs.
format_map(mapping)	Return a formatted version of S, using substitutions from mapping.
maketrans(x[, y, z])	Return a translation table usable for str.translate().

Attributes

INPUT
input
OUTPUT
output

capitalize(*/*)

Return a capitalized version of the string.

More specifically, make the first character have upper case and the rest lower case.

casefold(*/*)

Return a version of the string suitable for caseless comparisons.

center(*width*, *fillchar=' '*, */*)

Return a centered string of length width.

Padding is done using the specified fill character (default is a space).

count(*sub*[, *start*[, *end*]]) → int

Return the number of non-overlapping occurrences of substring sub in string S[start:end]. Optional arguments start and end are interpreted as in slice notation.

encode(/, *encoding='utf-8'*, *errors='strict'*)

Encode the string using the codec registered for encoding.

encoding
The encoding in which to encode the string.

errors
The error handling scheme to use for encoding errors. The default is 'strict' meaning that encoding errors raise a UnicodeEncodeError. Other possible values are 'ignore', 'replace' and 'xmlcharrefreplace' as well as any other name registered with codecs.register_error that can handle UnicodeEncodeErrors.

endswith(*suffix*[, *start*[, *end*]]) → bool

Return True if S ends with the specified suffix, False otherwise. With optional start, test S beginning at that position. With optional end, stop comparing S at that position. suffix can also be a tuple of strings to try.

expandtabs(/, *tabsize=8*)

Return a copy where all tab characters are expanded using spaces.

If tabsize is not given, a tab size of 8 characters is assumed.

find(*sub*[, *start*[, *end*]]) → int

Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end]. Optional arguments start and end are interpreted as in slice notation.

Return -1 on failure.

format(**args*, ***kwargs*) → str

Return a formatted version of S, using substitutions from args and kwargs. The substitutions are identified by braces ('{' and '}').

format_map(*mapping*) → str

Return a formatted version of S, using substitutions from mapping. The substitutions are identified by braces ('{' and '}').

index(*sub*[, *start*[, *end*]]) → int

Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end]. Optional arguments start and end are interpreted as in slice notation.

Raises ValueError when the substring is not found.

isalnum(/)

Return True if the string is an alpha-numeric string, False otherwise.

A string is alpha-numeric if all characters in the string are alpha-numeric and there is at least one character in the string.

isalpha(/)

Return True if the string is an alphabetic string, False otherwise.

A string is alphabetic if all characters in the string are alphabetic and there is at least one character in the string.

isascii(/)

Return True if all characters in the string are ASCII, False otherwise.

ASCII characters have code points in the range U+0000-U+007F. Empty string is ASCII too.

isdecimal(/)

Return True if the string is a decimal string, False otherwise.

A string is a decimal string if all characters in the string are decimal and there is at least one character in the string.

isdigit(/)

Return True if the string is a digit string, False otherwise.

A string is a digit string if all characters in the string are digits and there is at least one character in the string.

isidentifier(/)

Return True if the string is a valid Python identifier, False otherwise.

Call keyword.iskeyword(s) to test whether string s is a reserved identifier, such as "def" or "class".

islower(/)

Return True if the string is a lowercase string, False otherwise.

A string is lowercase if all cased characters in the string are lowercase and there is at least one cased character in the string.

isnumeric(/)

Return True if the string is a numeric string, False otherwise.

A string is numeric if all characters in the string are numeric and there is at least one character in the string.

isprintable(/)

Return True if the string is printable, False otherwise.

A string is printable if all of its characters are considered printable in repr() or if it is empty.

isspace(/)

Return True if the string is a whitespace string, False otherwise.

A string is whitespace if all characters in the string are whitespace and there is at least one character in the string.

istitle(/)

Return True if the string is a title-cased string, False otherwise.

In a title-cased string, upper- and title-case characters may only follow uncased characters and lowercase characters only cased ones.

isupper(/)

Return True if the string is an uppercase string, False otherwise.

A string is uppercase if all cased characters in the string are uppercase and there is at least one cased character in the string.

join(*iterable*, /)

Concatenate any number of strings.

The string whose method is called is inserted in between each given string. The result is returned as a new string.

Example: '.'.join(['ab', 'pq', 'rs']) -> 'ab.pq.rs'

ljust(*width*, *fillchar=' '*, */*)

Return a left-justified string of length width.

Padding is done using the specified fill character (default is a space).

lower(*/*)

Return a copy of the string converted to lowercase.

lstrip(*chars=None*, */*)

Return a copy of the string with leading whitespace removed.

If chars is given and not None, remove characters in chars instead.

static maketrans(*x*, *y=<unrepresentable>*, *z=<unrepresentable>*, */*)

Return a translation table usable for str.translate().

If there is only one argument, it must be a dictionary mapping Unicode ordinals (integers) or characters to Unicode ordinals, strings or None. Character keys will be then converted to ordinals. If there are two arguments, they must be strings of equal length, and in the resulting dictionary, each character in x will be mapped to the character at the same position in y. If there is a third argument, it must be a string, whose characters will be mapped to None in the result.

partition(*sep*, */*)

Partition the string into three parts using the given separator.

This will search for the separator in the string. If the separator is found, returns a 3-tuple containing the part before the separator, the separator itself, and the part after it.

If the separator is not found, returns a 3-tuple containing the original string and two empty strings.

removeprefix(*prefix*, */*)

Return a str with the given prefix string removed if present.

If the string starts with the prefix string, return string[len(prefix):]. Otherwise, return a copy of the original string.

removesuffix(*suffix*, */*)

Return a str with the given suffix string removed if present.

If the string ends with the suffix string and that suffix is not empty, return string[:-len(suffix)]. Otherwise, return a copy of the original string.

replace(*old*, *new*, *count=-1*, */*)

Return a copy with all occurrences of substring old replaced by new.

count
Maximum number of occurrences to replace. -1 (the default value) means replace all occurrences.

If the optional argument count is given, only the first count occurrences are replaced.

rfind(*sub*[, *start*[, *end*]]) → int

Return the highest index in S where substring sub is found, such that sub is contained within S[start:end]. Optional arguments start and end are interpreted as in slice notation.

Return -1 on failure.

rindex(*sub*[, *start*[, *end*]]) → int

Return the highest index in S where substring sub is found, such that sub is contained within S[start:end]. Optional arguments start and end are interpreted as in slice notation.

Raises ValueError when the substring is not found.

rjust(*width*, *fillchar=' '*, */*)

Return a right-justified string of length width.

Padding is done using the specified fill character (default is a space).

rpartition(*sep*, */*)

Partition the string into three parts using the given separator.

This will search for the separator in the string, starting at the end. If the separator is found, returns a 3-tuple containing the part before the separator, the separator itself, and the part after it.

If the separator is not found, returns a 3-tuple containing two empty strings and the original string.

rsplit(*/*, *sep=None*, *maxsplit=-1*)

Return a list of the substrings in the string, using sep as the separator string.

sep
The separator used to split the string.

When set to None (the default value), will split on any whitespace character (including n r t f and spaces) and will discard empty strings from the result.

maxsplit
Maximum number of splits. -1 (the default value) means no limit.

Splitting starts at the end of the string and works to the front.

rstrip(*chars=None*, */*)

Return a copy of the string with trailing whitespace removed.

If chars is given and not None, remove characters in chars instead.

split(*/*, *sep=None*, *maxsplit=-1*)

Return a list of the substrings in the string, using sep as the separator string.

sep
The separator used to split the string.

When set to None (the default value), will split on any whitespace character (including n r t f and spaces) and will discard empty strings from the result.

maxsplit
Maximum number of splits. -1 (the default value) means no limit.

Splitting starts at the front of the string and works to the end.

Note, str.split() is mainly useful for data that has been intentionally delimited. With natural text that includes punctuation, consider using the regular expression module.

splitlines(*/*, *keepends=False*)

Return a list of the lines in the string, breaking at line boundaries.

Line breaks are not included in the resulting list unless keepends is given and true.

startswith(*prefix*[, *start*[, *end*]]) → bool

Return True if S starts with the specified prefix, False otherwise. With optional start, test S beginning at that position. With optional end, stop comparing S at that position. prefix can also be a tuple of strings to try.

strip(*chars=None*, */*)

Return a copy of the string with leading and trailing whitespace removed.

If chars is given and not None, remove characters in chars instead.

swapcase(/)

Convert uppercase characters to lowercase and lowercase characters to uppercase.

title(/)

Return a version of the string where each word is titlecased.

More specifically, words start with uppercased characters and all remaining cased characters have lower case.

translate(*table*, /)

Replace each character in the string using the given translation table.

table
Translation table, which must be a mapping of Unicode ordinals to Unicode ordinals, strings, or None.

The table must implement lookup/indexing via __getitem__, for instance a dictionary or list. If this operation raises LookupError, the character is left untouched. Characters mapped to None are deleted.

upper(/)

Return a copy of the string converted to uppercase.

zfill(*width*, /)

Pad a numeric string with zeros on the left, to fill a field of the given width.

The string is never truncated.

libera_utils.constants.ProcessingStepIdentifier

class libera_utils.constants.**ProcessingStepIdentifier**(*value*)

Bases: StrEnum

Enumeration of processing step IDs used in AWS resource naming and processing orchestration.

In orchestration code, these are used as "NodeID" values to identify processing steps:
The processing_step_node_id values used in libera_cdk deployment of processing steps and the node names in processing_system_dag.json must match these.

They must also be passed to the ecr_upload module called by some libera_cdk integration tests.

The string values are the processing step names used in orchestration.

Each member is defined as a tuple: (step_name, products_list) - step_name: The string value used in orchestration and AWS resources - products_list: List of DataProductIdentifier members that this step produces

Example

```
>>> step = ProcessingStepIdentifier.l1b_rad
>>> str(step)  # Returns "l1b-rad"
>>> step.products  # Returns [DataProductIdentifier.l1b_rad]
>>> step.level  # Returns DataLevel.L1B (derived from products)
```

When adding new processing steps:

1. Add the enum member with its step name and list of produced products
2. No need to update any lookup dictionaries - relationships are embedded!

Methods

capitalize(/)	Return a capitalized version of the string.
casefold(/)	Return a version of the string suitable for caseless comparisons.
center(width[, fillchar])	Return a centered string of length width.
count(sub[, start[, end]])	Return the number of non-overlapping occurrences of substring sub in string S[start:end].
encode(/[, encoding, errors])	Encode the string using the codec registered for encoding.
endswith(suffix[, start[, end]])	Return True if S ends with the specified suffix, False otherwise.
expandtabs(/[, tabsize])	Return a copy where all tab characters are expanded using spaces.
find(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
format(*args, **kwargs)	Return a formatted version of S, using substitutions from args and kwargs.
format_map(mapping)	Return a formatted version of S, using substitutions from mapping.
index(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
isalnum(/)	Return True if the string is an alpha-numeric string, False otherwise.
isalpha(/)	Return True if the string is an alphabetic string, False otherwise.
isascii(/)	Return True if all characters in the string are ASCII, False otherwise.
isdecimal(/)	Return True if the string is a decimal string, False otherwise.
isdigit(/)	Return True if the string is a digit string, False otherwise.
isidentifier(/)	Return True if the string is a valid Python identifier, False otherwise.
islower(/)	Return True if the string is a lowercase string, False otherwise.
isnumeric(/)	Return True if the string is a numeric string, False otherwise.
isprintable(/)	Return True if the string is printable, False otherwise.
isspace(/)	Return True if the string is a whitespace string, False otherwise.
istitle(/)	Return True if the string is a title-cased string, False otherwise.
isupper(/)	Return True if the string is an uppercase string, False otherwise.
join(iterable, /)	Concatenate any number of strings.
ljust(width[, fillchar])	Return a left-justified string of length width.
lower(/)	Return a copy of the string converted to lowercase.
lstrip([chars])	Return a copy of the string with leading whitespace removed.
maketrans(x[, y, z])	Return a translation table usable for str.translate().
partition(sep, /)	Partition the string into three parts using the given separator.

continues on next page

Table 41 – continued from previous page

removeprefix(prefix, /)	Return a str with the given prefix string removed if present.
removesuffix(suffix, /)	Return a str with the given suffix string removed if present.
replace(old, new[, count])	Return a copy with all occurrences of substring old replaced by new.
rfind(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rindex(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rjust(width[, fillchar])	Return a right-justified string of length width.
rpartition(sep, /)	Partition the string into three parts using the given separator.
rsplit(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
rstrip([chars])	Return a copy of the string with trailing whitespace removed.
split(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
splitlines(/[, keepends])	Return a list of the lines in the string, breaking at line boundaries.
startswith(prefix[, start[, end]])	Return True if S starts with the specified prefix, False otherwise.
strip([chars])	Return a copy of the string with leading and trailing whitespace removed.
swapcase(/)	Convert uppercase characters to lowercase and lowercase characters to uppercase.
title(/)	Return a version of the string where each word is titlecased.
translate(table, /)	Replace each character in the string using the given translation table.
upper(/)	Return a copy of the string converted to uppercase.
zfill(width, /)	Pad a numeric string with zeros on the left, to fill a field of the given width.

__init__(**args*, ***kwds*)

Methods

encode(/[, encoding, errors])	Encode the string using the codec registered for encoding.
replace(old, new[, count])	Return a copy with all occurrences of substring old replaced by new.
split(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
rsplit(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
join(iterable, /)	Concatenate any number of strings.
capitalize(/)	Return a capitalized version of the string.

continues on next page

Table 42 – continued from previous page

casefold(/)	Return a version of the string suitable for caseless comparisons.
title(/)	Return a version of the string where each word is titlecased.
center(width[, fillchar])	Return a centered string of length width.
count(sub[, start[, end]])	Return the number of non-overlapping occurrences of substring sub in string S[start:end].
expandtabs(/[, tabsize])	Return a copy where all tab characters are expanded using spaces.
find(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
partition(sep, /)	Partition the string into three parts using the given separator.
index(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
ljust(width[, fillchar])	Return a left-justified string of length width.
lower(/)	Return a copy of the string converted to lowercase.
lstrip([chars])	Return a copy of the string with leading whitespace removed.
rfind(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rindex(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rjust(width[, fillchar])	Return a right-justified string of length width.
rstrip([chars])	Return a copy of the string with trailing whitespace removed.
rpartition(sep, /)	Partition the string into three parts using the given separator.
splitlines(/[, keepends])	Return a list of the lines in the string, breaking at line boundaries.
strip([chars])	Return a copy of the string with leading and trailing whitespace removed.
swapcase(/)	Convert uppercase characters to lowercase and lowercase characters to uppercase.
translate(table, /)	Replace each character in the string using the given translation table.
upper(/)	Return a copy of the string converted to uppercase.
startswith(prefix[, start[, end]])	Return True if S starts with the specified prefix, False otherwise.
endswith(suffix[, start[, end]])	Return True if S ends with the specified suffix, False otherwise.
removeprefix(prefix, /)	Return a str with the given prefix string removed if present.
removesuffix(suffix, /)	Return a str with the given suffix string removed if present.
isascii(/)	Return True if all characters in the string are ASCII, False otherwise.
islower(/)	Return True if the string is a lowercase string, False otherwise.
isupper(/)	Return True if the string is an uppercase string, False otherwise.

continues on next page

Table 42 – continued from previous page

istitle(/)	Return True if the string is a title-cased string, False otherwise.
isspace(/)	Return True if the string is a whitespace string, False otherwise.
isdecimal(/)	Return True if the string is a decimal string, False otherwise.
isdigit(/)	Return True if the string is a digit string, False otherwise.
isnumeric(/)	Return True if the string is a numeric string, False otherwise.
isalpha(/)	Return True if the string is an alphabetic string, False otherwise.
isalnum(/)	Return True if the string is an alpha-numeric string, False otherwise.
isidentifier(/)	Return True if the string is a valid Python identifier, False otherwise.
isprintable(/)	Return True if the string is printable, False otherwise.
zfill(width, /)	Pad a numeric string with zeros on the left, to fill a field of the given width.
format(*args, **kwargs)	Return a formatted version of S, using substitutions from args and kwargs.
format_map(mapping)	Return a formatted version of S, using substitutions from mapping.
maketrans(x[, y, z])	Return a translation table usable for str.translate().
get_archive_bucket_name([account_suffix])	Gets the archive bucket name for this processing step.
from_data_product(data_product)	Get the ProcessingStepIdentifier that produces the given DataProductIdentifier

Attributes

processing_step_name	Get the name formatted for AWS resources for this processing step
products	Get the list of data products produced by this processing step.
level	Get the processing level of the products produced by this step
step_function_name	Get the name formatted for the step function for this processing step
policy_name	Get the name formatted IAM policy for this processing step
ecr_name	Get the manually-configured ECR name for this processing step
`l1a_rad`	
`l1a_cam`	
`l1a_jpss`	
`l1a_azel`	

continues on next page

Table 43 – continued from previous page

cal_rad
cal_cam
spice_azel
spice_jpss
l1b_rad
l1b_cam
int_footprint_scene_id
l2_cf_rad
l2_cf_cam
l2_unfiltered
l2_ssw_toa_osse
l2_ssw_toa_erbe
l2_ssw_toa_trmm
l2_ssw_toa_rt
l2_surface_flux
adm_binning

capitalize(/)

Return a capitalized version of the string.

More specifically, make the first character have upper case and the rest lower case.

casefold(/)

Return a version of the string suitable for caseless comparisons.

center(*width*, *fillchar=' '*, /)

Return a centered string of length width.

Padding is done using the specified fill character (default is a space).

count(*sub*[, *start*[, *end*]]) → int

Return the number of non-overlapping occurrences of substring sub in string S[start:end]. Optional arguments start and end are interpreted as in slice notation.

property ecr_name: str | None

Get the manually-configured ECR name for this processing step

We name our ECRs in CDK because they are one of the few resources that humans will need to interact with on a regular basis.

encode(*/*, *encoding='utf-8'*, *errors='strict'*)

Encode the string using the codec registered for encoding.

encoding
The encoding in which to encode the string.

errors
The error handling scheme to use for encoding errors. The default is 'strict' meaning that encoding errors raise a UnicodeEncodeError. Other possible values are 'ignore', 'replace' and 'xmlcharrefreplace' as well as any other name registered with codecs.register_error that can handle UnicodeEncodeErrors.

endswith(*suffix*[, *start*[, *end*]]) → bool

Return True if S ends with the specified suffix, False otherwise. With optional start, test S beginning at that position. With optional end, stop comparing S at that position. suffix can also be a tuple of strings to try.

expandtabs(*/*, *tabsize=8*)

Return a copy where all tab characters are expanded using spaces.

If tabsize is not given, a tab size of 8 characters is assumed.

find(*sub*[, *start*[, *end*]]) → int

Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end]. Optional arguments start and end are interpreted as in slice notation.

Return -1 on failure.

format(**args*, ***kwargs*) → str

Return a formatted version of S, using substitutions from args and kwargs. The substitutions are identified by braces ('{' and '}').

format_map(*mapping*) → str

Return a formatted version of S, using substitutions from mapping. The substitutions are identified by braces ('{' and '}').

classmethod from_data_product(*data_product:* DataProductIdentifier) → *ProcessingStepIdentifier* | None

Get the ProcessingStepIdentifier that produces the given DataProductIdentifier

Parameters
data_product (DataProductIdentifier) – The data product to find the processing step for

Returns
The processing step that produces this data product

Return type
ProcessingStepIdentifier

Raises
ValueError – If no processing step is found for the data product

get_archive_bucket_name(*account_suffix: str* | LiberaAccountSuffix = *LiberaAccountSuffix.STAGE*) → str | None

Gets the archive bucket name for this processing step.

Buckets are named according to the level of data they are storing and which account they are in. This is expected to be used by the L2 developers who will most commonly be working with the stage account.

Parameters
account_suffix (*str | LiberaAccountSuffix, optional*) – Account suffix for the

bucket name, by default LiberaAccountSuffix.STAGE (stage account). Can be a string like "-test" for custom testing scenarios.

Returns
The name of the archive bucket for this processing step

Return type
str

index(*sub*[, *start*[, *end*]]) → int

Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end]. Optional arguments start and end are interpreted as in slice notation.

Raises ValueError when the substring is not found.

isalnum()

Return True if the string is an alpha-numeric string, False otherwise.

A string is alpha-numeric if all characters in the string are alpha-numeric and there is at least one character in the string.

isalpha()

Return True if the string is an alphabetic string, False otherwise.

A string is alphabetic if all characters in the string are alphabetic and there is at least one character in the string.

isascii()

Return True if all characters in the string are ASCII, False otherwise.

ASCII characters have code points in the range U+0000-U+007F. Empty string is ASCII too.

isdecimal()

Return True if the string is a decimal string, False otherwise.

A string is a decimal string if all characters in the string are decimal and there is at least one character in the string.

isdigit()

Return True if the string is a digit string, False otherwise.

A string is a digit string if all characters in the string are digits and there is at least one character in the string.

isidentifier()

Return True if the string is a valid Python identifier, False otherwise.

Call keyword.iskeyword(s) to test whether string s is a reserved identifier, such as "def" or "class".

islower()

Return True if the string is a lowercase string, False otherwise.

A string is lowercase if all cased characters in the string are lowercase and there is at least one cased character in the string.

isnumeric()

Return True if the string is a numeric string, False otherwise.

A string is numeric if all characters in the string are numeric and there is at least one character in the string.

isprintable(/)

Return True if the string is printable, False otherwise.

A string is printable if all of its characters are considered printable in repr() or if it is empty.

isspace(/)

Return True if the string is a whitespace string, False otherwise.

A string is whitespace if all characters in the string are whitespace and there is at least one character in the string.

istitle(/)

Return True if the string is a title-cased string, False otherwise.

In a title-cased string, upper- and title-case characters may only follow uncased characters and lowercase characters only cased ones.

isupper(/)

Return True if the string is an uppercase string, False otherwise.

A string is uppercase if all cased characters in the string are uppercase and there is at least one cased character in the string.

join(*iterable*, /)

Concatenate any number of strings.

The string whose method is called is inserted in between each given string. The result is returned as a new string.

Example: '.'.join(['ab', 'pq', 'rs']) -> 'ab.pq.rs'

property level: *DataLevel*

Get the processing level of the products produced by this step

Raises
ValueError – If the step produces no products or produces products of multiple levels

ljust(*width*, *fillchar=' '*, /)

Return a left-justified string of length width.

Padding is done using the specified fill character (default is a space).

lower(/)

Return a copy of the string converted to lowercase.

lstrip(*chars=None*, /)

Return a copy of the string with leading whitespace removed.

If chars is given and not None, remove characters in chars instead.

static maketrans(*x*, *y=<unrepresentable>*, *z=<unrepresentable>*, /)

Return a translation table usable for str.translate().

If there is only one argument, it must be a dictionary mapping Unicode ordinals (integers) or characters to Unicode ordinals, strings or None. Character keys will be then converted to ordinals. If there are two arguments, they must be strings of equal length, and in the resulting dictionary, each character in x will be mapped to the character at the same position in y. If there is a third argument, it must be a string, whose characters will be mapped to None in the result.

partition(*sep*, /)

Partition the string into three parts using the given separator.

This will search for the separator in the string. If the separator is found, returns a 3-tuple containing the part before the separator, the separator itself, and the part after it.

If the separator is not found, returns a 3-tuple containing the original string and two empty strings.

property policy_name: str

Get the name formatted IAM policy for this processing step

The policy name is used to create an IAM policy that grants permissions to the aspects of the processing step.

property processing_step_name: str

Get the name formatted for AWS resources for this processing step

The name is used to create AWS resources that are specific to the processing step. This is an alias to the string value for compatibility.

property products: list[*DataProductIdentifier*]

Get the list of data products produced by this processing step.

Returns
List of data products produced by this processing step

Return type
list[*DataProductIdentifier*]

removeprefix(*prefix*, /)

Return a str with the given prefix string removed if present.

If the string starts with the prefix string, return string[len(prefix):]. Otherwise, return a copy of the original string.

removesuffix(*suffix*, /)

Return a str with the given suffix string removed if present.

If the string ends with the suffix string and that suffix is not empty, return string[:-len(suffix)]. Otherwise, return a copy of the original string.

replace(*old*, *new*, *count=-1*, /)

Return a copy with all occurrences of substring old replaced by new.

count
Maximum number of occurrences to replace. -1 (the default value) means replace all occurrences.

If the optional argument count is given, only the first count occurrences are replaced.

rfind(*sub*[, *start*[, *end*]]) → int

Return the highest index in S where substring sub is found, such that sub is contained within S[start:end]. Optional arguments start and end are interpreted as in slice notation.

Return -1 on failure.

rindex(*sub*[, *start*[, *end*]]) → int

Return the highest index in S where substring sub is found, such that sub is contained within S[start:end]. Optional arguments start and end are interpreted as in slice notation.

Raises ValueError when the substring is not found.

rjust(*width*, *fillchar=' '*, */*)

Return a right-justified string of length width.

Padding is done using the specified fill character (default is a space).

rpartition(*sep*, */*)

Partition the string into three parts using the given separator.

This will search for the separator in the string, starting at the end. If the separator is found, returns a 3-tuple containing the part before the separator, the separator itself, and the part after it.

If the separator is not found, returns a 3-tuple containing two empty strings and the original string.

rsplit(*/*, *sep=None*, *maxsplit=-1*)

Return a list of the substrings in the string, using sep as the separator string.

sep
The separator used to split the string.

When set to None (the default value), will split on any whitespace character (including n r t f and spaces) and will discard empty strings from the result.

maxsplit
Maximum number of splits. -1 (the default value) means no limit.

Splitting starts at the end of the string and works to the front.

rstrip(*chars=None*, */*)

Return a copy of the string with trailing whitespace removed.

If chars is given and not None, remove characters in chars instead.

split(*/*, *sep=None*, *maxsplit=-1*)

Return a list of the substrings in the string, using sep as the separator string.

sep
The separator used to split the string.

When set to None (the default value), will split on any whitespace character (including n r t f and spaces) and will discard empty strings from the result.

maxsplit
Maximum number of splits. -1 (the default value) means no limit.

Splitting starts at the front of the string and works to the end.

Note, str.split() is mainly useful for data that has been intentionally delimited. With natural text that includes punctuation, consider using the regular expression module.

splitlines(*/*, *keepends=False*)

Return a list of the lines in the string, breaking at line boundaries.

Line breaks are not included in the resulting list unless keepends is given and true.

startswith(*prefix*[, *start*[, *end*]]) → bool

Return True if S starts with the specified prefix, False otherwise. With optional start, test S beginning at that position. With optional end, stop comparing S at that position. prefix can also be a tuple of strings to try.

property step_function_name

Get the name formatted for the step function for this processing step

The step function name is used to create a step function that orchestrates the processing step.

strip(*chars=None*, /)

Return a copy of the string with leading and trailing whitespace removed.

If chars is given and not None, remove characters in chars instead.

swapcase(/)

Convert uppercase characters to lowercase and lowercase characters to uppercase.

title(/)

Return a version of the string where each word is titlecased.

More specifically, words start with uppercased characters and all remaining cased characters have lower case.

translate(*table*, /)

Replace each character in the string using the given translation table.

table
Translation table, which must be a mapping of Unicode ordinals to Unicode ordinals, strings, or None.

The table must implement lookup/indexing via __getitem__, for instance a dictionary or list. If this operation raises LookupError, the character is left untouched. Characters mapped to None are deleted.

upper(/)

Return a copy of the string converted to uppercase.

zfill(*width*, /)

Pad a numeric string with zeros on the left, to fill a field of the given width.

The string is never truncated.

class libera_utils.constants.**DataLevel**(*value*)

Data product level

Methods

capitalize(/)	Return a capitalized version of the string.
casefold(/)	Return a version of the string suitable for caseless comparisons.
center(width[, fillchar])	Return a centered string of length width.
count(sub[, start[, end]])	Return the number of non-overlapping occurrences of substring sub in string S[start:end].
encode(/[, encoding, errors])	Encode the string using the codec registered for encoding.
endswith(suffix[, start[, end]])	Return True if S ends with the specified suffix, False otherwise.
expandtabs(/[, tabsize])	Return a copy where all tab characters are expanded using spaces.
find(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
format(*args, **kwargs)	Return a formatted version of S, using substitutions from args and kwargs.
format_map(mapping)	Return a formatted version of S, using substitutions from mapping.

continues on next page

Table 44 – continued from previous page

index(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
isalnum(/)	Return True if the string is an alpha-numeric string, False otherwise.
isalpha(/)	Return True if the string is an alphabetic string, False otherwise.
isascii(/)	Return True if all characters in the string are ASCII, False otherwise.
isdecimal(/)	Return True if the string is a decimal string, False otherwise.
isdigit(/)	Return True if the string is a digit string, False otherwise.
isidentifier(/)	Return True if the string is a valid Python identifier, False otherwise.
islower(/)	Return True if the string is a lowercase string, False otherwise.
isnumeric(/)	Return True if the string is a numeric string, False otherwise.
isprintable(/)	Return True if the string is printable, False otherwise.
isspace(/)	Return True if the string is a whitespace string, False otherwise.
istitle(/)	Return True if the string is a title-cased string, False otherwise.
isupper(/)	Return True if the string is an uppercase string, False otherwise.
join(iterable, /)	Concatenate any number of strings.
ljust(width[, fillchar])	Return a left-justified string of length width.
lower(/)	Return a copy of the string converted to lowercase.
lstrip([chars])	Return a copy of the string with leading whitespace removed.
maketrans(x[, y, z])	Return a translation table usable for str.translate().
partition(sep, /)	Partition the string into three parts using the given separator.
removeprefix(prefix, /)	Return a str with the given prefix string removed if present.
removesuffix(suffix, /)	Return a str with the given suffix string removed if present.
replace(old, new[, count])	Return a copy with all occurrences of substring old replaced by new.
rfind(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rindex(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rjust(width[, fillchar])	Return a right-justified string of length width.
rpartition(sep, /)	Partition the string into three parts using the given separator.
rsplit(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
rstrip([chars])	Return a copy of the string with trailing whitespace removed.
split(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.

continues on next page

Table 44 – continued from previous page

splitlines(/[, keepends])	Return a list of the lines in the string, breaking at line boundaries.
startswith(prefix[, start[, end]])	Return True if S starts with the specified prefix, False otherwise.
strip([chars])	Return a copy of the string with leading and trailing whitespace removed.
swapcase(/)	Convert uppercase characters to lowercase and lowercase characters to uppercase.
title(/)	Return a version of the string where each word is titlecased.
translate(table, /)	Replace each character in the string using the given translation table.
upper(/)	Return a copy of the string converted to uppercase.
zfill(width, /)	Pad a numeric string with zeros on the left, to fill a field of the given width.

property archive_bucket_name: str

Gets the archive bucket name for the data level.

Notes

This does not include any account suffix, which must be added separately.

class libera_utils.constants.**DataProductIdentifier**(*value*)

Enumeration of data product canonical IDs used in AWS resource naming.

These IDs refer to the data products (files) themselves, NOT the processing steps (since processing steps may produce multiple products). The string values are the product names used in filenames.

This enum replaces the old ProductName enum from filenaming.py to provide a single source of truth.

Each member is defined as a tuple: (product_name, data_level) - product_name: The string value used in filenames and AWS resources - data_level: The DataLevel enum value indicating the processing level

Example

```
>>> product = DataProductIdentifier.l1b_rad
>>> str(product)  # Returns "RAD-4CH"
>>> product.level  # Returns DataLevel.L1B
>>> product.level.archive_bucket_name  # Returns "libera-l1b-data"
```

When adding new products:

1. Add the enum member with its product name and DataLevel
2. No need to update any lookup dictionaries - metadata is embedded!

Methods

capitalize(/)	Return a capitalized version of the string.
casefold(/)	Return a version of the string suitable for caseless comparisons.
center(width[, fillchar])	Return a centered string of length width.

continues on next page

Table 45 – continued from previous page

count(sub[, start[, end]])	Return the number of non-overlapping occurrences of substring sub in string S[start:end].
encode(/[, encoding, errors])	Encode the string using the codec registered for encoding.
endswith(suffix[, start[, end]])	Return True if S ends with the specified suffix, False otherwise.
expandtabs(/[, tabsize])	Return a copy where all tab characters are expanded using spaces.
find(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
format(*args, **kwargs)	Return a formatted version of S, using substitutions from args and kwargs.
format_map(mapping)	Return a formatted version of S, using substitutions from mapping.
index(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
isalnum(/)	Return True if the string is an alpha-numeric string, False otherwise.
isalpha(/)	Return True if the string is an alphabetic string, False otherwise.
isascii(/)	Return True if all characters in the string are ASCII, False otherwise.
isdecimal(/)	Return True if the string is a decimal string, False otherwise.
isdigit(/)	Return True if the string is a digit string, False otherwise.
isidentifier(/)	Return True if the string is a valid Python identifier, False otherwise.
islower(/)	Return True if the string is a lowercase string, False otherwise.
isnumeric(/)	Return True if the string is a numeric string, False otherwise.
isprintable(/)	Return True if the string is printable, False otherwise.
isspace(/)	Return True if the string is a whitespace string, False otherwise.
istitle(/)	Return True if the string is a title-cased string, False otherwise.
isupper(/)	Return True if the string is an uppercase string, False otherwise.
join(iterable, /)	Concatenate any number of strings.
ljust(width[, fillchar])	Return a left-justified string of length width.
lower(/)	Return a copy of the string converted to lowercase.
lstrip([chars])	Return a copy of the string with leading whitespace removed.
maketrans(x[, y, z])	Return a translation table usable for str.translate().
partition(sep, /)	Partition the string into three parts using the given separator.
removeprefix(prefix, /)	Return a str with the given prefix string removed if present.
removesuffix(suffix, /)	Return a str with the given suffix string removed if present.

continues on next page

Table 45 – continued from previous page

replace(old, new[, count])	Return a copy with all occurrences of substring old replaced by new.
rfind(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rindex(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rjust(width[, fillchar])	Return a right-justified string of length width.
rpartition(sep, /)	Partition the string into three parts using the given separator.
rsplit(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
rstrip([chars])	Return a copy of the string with trailing whitespace removed.
split(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
splitlines(/[, keepends])	Return a list of the lines in the string, breaking at line boundaries.
startswith(prefix[, start[, end]])	Return True if S starts with the specified prefix, False otherwise.
strip([chars])	Return a copy of the string with leading and trailing whitespace removed.
swapcase(/)	Convert uppercase characters to lowercase and lowercase characters to uppercase.
title(/)	Return a version of the string where each word is titlecased.
translate(table, /)	Replace each character in the string using the given translation table.
upper(/)	Return a copy of the string converted to uppercase.
zfill(width, /)	Pad a numeric string with zeros on the left, to fill a field of the given width.

property data_level: *DataLevel*

Get the processing level for this data product.

Returns
The processing level of this data product

Return type
DataLevel

get_partial_archive_bucket_name() → str

Gets the archive bucket name from the data product identifier .

Buckets are named according to the level of data they are storing and which account they are in. This is expected to be used by the L2 developers who will most commonly be working with the stage account.

Returns
The name of the archive bucket for this data product without an account suffix

Return type
str

property product_name: str

Get the name formatted for AWS resources for this data product

The name is used to create AWS resources that are specific to the data product. This is an alias to the string value for compatibility.

class libera_utils.constants.**LiberaApid**(*value*)

APIDs for packets

The enum names here should be of the form <packet-source>_<system>_<contents>. e.g. for icie_rad_sample: packet_source is Libera "ICIE", the system is the "Radiometers", and the contents is radiometer "Samples".

Notes

This is useful for identifying the data product type from the APID in an L0 filename.

The enum names (e.g. icie_seq_hk) for Libera packets here should be precisely the packet names used in FSW documents and packet definitions.

Attributes

denominator
the denominator of a rational number in lowest terms

imag
the imaginary part of a complex number

numerator
the numerator of a rational number in lowest terms

real
the real part of a complex number

Methods

as_integer_ratio(/)	Return integer ratio.
bit_count(/)	Number of ones in the binary representation of the absolute value of self.
bit_length(/)	Number of bits necessary to represent self in binary.
conjugate	Returns self, the complex conjugate of any int.
from_bytes(/, bytes[, byteorder, signed])	Return the integer represented by the given array of bytes.
to_bytes(/[, length, byteorder, signed])	Return an array of bytes representing an integer.

property **data_product_id**: *DataProductIdentifier*

Get the DataProductIdentifier for L0 PDS files with this APID

class libera_utils.constants.**ManifestType**(*value*)

Enumerated legal manifest type values

Methods

capitalize(/)	Return a capitalized version of the string.
casefold(/)	Return a version of the string suitable for caseless comparisons.
center(width[, fillchar])	Return a centered string of length width.
count(sub[, start[, end]])	Return the number of non-overlapping occurrences of substring sub in string S[start:end].

continues on next page

Table 47 – continued from previous page

encode(/[, encoding, errors])	Encode the string using the codec registered for encoding.
endswith(suffix[, start[, end]])	Return True if S ends with the specified suffix, False otherwise.
expandtabs(/[, tabsize])	Return a copy where all tab characters are expanded using spaces.
find(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
format(*args, **kwargs)	Return a formatted version of S, using substitutions from args and kwargs.
format_map(mapping)	Return a formatted version of S, using substitutions from mapping.
index(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
isalnum(/)	Return True if the string is an alpha-numeric string, False otherwise.
isalpha(/)	Return True if the string is an alphabetic string, False otherwise.
isascii(/)	Return True if all characters in the string are ASCII, False otherwise.
isdecimal(/)	Return True if the string is a decimal string, False otherwise.
isdigit(/)	Return True if the string is a digit string, False otherwise.
isidentifier(/)	Return True if the string is a valid Python identifier, False otherwise.
islower(/)	Return True if the string is a lowercase string, False otherwise.
isnumeric(/)	Return True if the string is a numeric string, False otherwise.
isprintable(/)	Return True if the string is printable, False otherwise.
isspace(/)	Return True if the string is a whitespace string, False otherwise.
istitle(/)	Return True if the string is a title-cased string, False otherwise.
isupper(/)	Return True if the string is an uppercase string, False otherwise.
join(iterable, /)	Concatenate any number of strings.
ljust(width[, fillchar])	Return a left-justified string of length width.
lower(/)	Return a copy of the string converted to lowercase.
lstrip([chars])	Return a copy of the string with leading whitespace removed.
maketrans(x[, y, z])	Return a translation table usable for str.translate().
partition(sep, /)	Partition the string into three parts using the given separator.
removeprefix(prefix, /)	Return a str with the given prefix string removed if present.
removesuffix(suffix, /)	Return a str with the given suffix string removed if present.
replace(old, new[, count])	Return a copy with all occurrences of substring old replaced by new.

continues on next page

Table 47 – continued from previous page

rfind(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rindex(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rjust(width[, fillchar])	Return a right-justified string of length width.
rpartition(sep, /)	Partition the string into three parts using the given separator.
rsplit(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
rstrip([chars])	Return a copy of the string with trailing whitespace removed.
split(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
splitlines(/[, keepends])	Return a list of the lines in the string, breaking at line boundaries.
startswith(prefix[, start[, end]])	Return True if S starts with the specified prefix, False otherwise.
strip([chars])	Return a copy of the string with leading and trailing whitespace removed.
swapcase(/)	Convert uppercase characters to lowercase and lowercase characters to uppercase.
title(/)	Return a version of the string where each word is titlecased.
translate(table, /)	Replace each character in the string using the given translation table.
upper(/)	Return a copy of the string converted to uppercase.
zfill(width, /)	Pad a numeric string with zeros on the left, to fill a field of the given width.

class libera_utils.constants.**ProcessingStepIdentifier**(*value*)

Enumeration of processing step IDs used in AWS resource naming and processing orchestration.

In orchestration code, these are used as "NodeID" values to identify processing steps:

The processing_step_node_id values used in libera_cdk deployment of processing steps and the node names in processing_system_dag.json must match these.

They must also be passed to the ecr_upload module called by some libera_cdk integration tests.

The string values are the processing step names used in orchestration.

Each member is defined as a tuple: (step_name, products_list) - step_name: The string value used in orchestration and AWS resources - products_list: List of DataProductIdentifier members that this step produces

Example

```
>>> step = ProcessingStepIdentifier.l1b_rad
>>> str(step)  # Returns "l1b-rad"
>>> step.products  # Returns [DataProductIdentifier.l1b_rad]
>>> step.level  # Returns DataLevel.L1B (derived from products)
```

When adding new processing steps:

1. Add the enum member with its step name and list of produced products
2. No need to update any lookup dictionaries - relationships are embedded!

Methods

capitalize(/)	Return a capitalized version of the string.
casefold(/)	Return a version of the string suitable for caseless comparisons.
center(width[, fillchar])	Return a centered string of length width.
count(sub[, start[, end]])	Return the number of non-overlapping occurrences of substring sub in string S[start:end].
encode(/[, encoding, errors])	Encode the string using the codec registered for encoding.
endswith(suffix[, start[, end]])	Return True if S ends with the specified suffix, False otherwise.
expandtabs(/[, tabsize])	Return a copy where all tab characters are expanded using spaces.
find(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
format(*args, **kwargs)	Return a formatted version of S, using substitutions from args and kwargs.
format_map(mapping)	Return a formatted version of S, using substitutions from mapping.
index(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
isalnum(/)	Return True if the string is an alpha-numeric string, False otherwise.
isalpha(/)	Return True if the string is an alphabetic string, False otherwise.
isascii(/)	Return True if all characters in the string are ASCII, False otherwise.
isdecimal(/)	Return True if the string is a decimal string, False otherwise.
isdigit(/)	Return True if the string is a digit string, False otherwise.
isidentifier(/)	Return True if the string is a valid Python identifier, False otherwise.
islower(/)	Return True if the string is a lowercase string, False otherwise.
isnumeric(/)	Return True if the string is a numeric string, False otherwise.
isprintable(/)	Return True if the string is printable, False otherwise.
isspace(/)	Return True if the string is a whitespace string, False otherwise.
istitle(/)	Return True if the string is a title-cased string, False otherwise.
isupper(/)	Return True if the string is an uppercase string, False otherwise.
join(iterable, /)	Concatenate any number of strings.
ljust(width[, fillchar])	Return a left-justified string of length width.
lower(/)	Return a copy of the string converted to lowercase.
lstrip([chars])	Return a copy of the string with leading whitespace removed.
maketrans(x[, y, z])	Return a translation table usable for str.translate().
partition(sep, /)	Partition the string into three parts using the given separator.

continues on next page

Table 48 – continued from previous page

removeprefix(prefix, /)	Return a str with the given prefix string removed if present.
removesuffix(suffix, /)	Return a str with the given suffix string removed if present.
replace(old, new[, count])	Return a copy with all occurrences of substring old replaced by new.
rfind(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rindex(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rjust(width[, fillchar])	Return a right-justified string of length width.
rpartition(sep, /)	Partition the string into three parts using the given separator.
rsplit(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
rstrip([chars])	Return a copy of the string with trailing whitespace removed.
split(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
splitlines(/[, keepends])	Return a list of the lines in the string, breaking at line boundaries.
startswith(prefix[, start[, end]])	Return True if S starts with the specified prefix, False otherwise.
strip([chars])	Return a copy of the string with leading and trailing whitespace removed.
swapcase(/)	Convert uppercase characters to lowercase and lowercase characters to uppercase.
title(/)	Return a version of the string where each word is titlecased.
translate(table, /)	Replace each character in the string using the given translation table.
upper(/)	Return a copy of the string converted to uppercase.
zfill(width, /)	Pad a numeric string with zeros on the left, to fill a field of the given width.

property ecr_name: `str | None`

Get the manually-configured ECR name for this processing step

We name our ECRs in CDK because they are one of the few resources that humans will need to interact with on a regular basis.

classmethod from_data_product(*data_product:* DataProductIdentifier) → *ProcessingStepIdentifier* | None

Get the ProcessingStepIdentifier that produces the given DataProductIdentifier

Parameters
data_product (DataProductIdentifier) – The data product to find the processing step for

Returns
The processing step that produces this data product

Return type
ProcessingStepIdentifier

Raises
ValueError – If no processing step is found for the data product

get_archive_bucket_name(*account_suffix: str | LiberaAccountSuffix = LiberaAccountSuffix.STAGE*) → str | None

Gets the archive bucket name for this processing step.

Buckets are named according to the level of data they are storing and which account they are in. This is expected to be used by the L2 developers who will most commonly be working with the stage account.

Parameters
account_suffix (*str | LiberaAccountSuffix, optional*) – Account suffix for the bucket name, by default LiberaAccountSuffix.STAGE (stage account). Can be a string like "-test" for custom testing scenarios.

Returns
The name of the archive bucket for this processing step

Return type
str

property level: *DataLevel*

Get the processing level of the products produced by this step

Raises
ValueError – If the step produces no products or produces products of multiple levels

property policy_name: str

Get the name formatted IAM policy for this processing step

The policy name is used to create an IAM policy that grants permissions to the aspects of the processing step.

property processing_step_name: str

Get the name formatted for AWS resources for this processing step

The name is used to create AWS resources that are specific to the processing step. This is an alias to the string value for compatibility.

property products: list[*DataProductIdentifier*]

Get the list of data products produced by this processing step.

Returns
List of data products produced by this processing step

Return type
list[*DataProductIdentifier*]

property step_function_name

Get the name formatted for the step function for this processing step

The step function name is used to create a step function that orchestrates the processing step.

3.1.5 libera_utils.db

db module for dyanmodb operations.

Modules

dynamodb_utils	Module for database utilities

libera_utils.db.dynamodb_utils

Module for database utilities

Functions

add_archive_time_to_ddb_item(ddb_item)	Add archive time to DynamoDB item
create_ddb_metadata_applicable_date_item(*, ...)	Write metadata record to DynamoDB for a single file
create_ddb_metadata_file_item(filename, ...)	Write metadata record to DynamoDB for a single file
get_dynamodb_table(dynamo_table_name)	Get the DynamoDB table

libera_utils.db.dynamodb_utils.add_archive_time_to_ddb_item

libera_utils.db.dynamodb_utils.**add_archive_time_to_ddb_item**(*ddb_item: dict*)

Add archive time to DynamoDB item

libera_utils.db.dynamodb_utils.create_ddb_metadata_applicable_date_item

libera_utils.db.dynamodb_utils.**create_ddb_metadata_applicable_date_item**(*, *filename: str*, *data_level: str*, *data_type: str*, *applicable_date: str*, *data_subtype: str = None*, *additional_metadata: dict = None*)

Write metadata record to DynamoDB for a single file

libera_utils.db.dynamodb_utils.create_ddb_metadata_file_item

libera_utils.db.dynamodb_utils.**create_ddb_metadata_file_item**(*filename: str*, *algorithm_version: str*, *include_archive_time: bool = False*, *additional_metadata: dict = None*)

Write metadata record to DynamoDB for a single file

libera_utils.db.dynamodb_utils.get_dynamodb_table

libera_utils.db.dynamodb_utils.**get_dynamodb_table**(*dynamo_table_name: str*)

Get the DynamoDB table

libera_utils.db.dynamodb_utils.**add_archive_time_to_ddb_item**(*ddb_item: dict*)

Add archive time to DynamoDB item

libera_utils.db.dynamodb_utils.**create_ddb_metadata_applicable_date_item**(*, *filename: str*, *data_level: str*, *data_type: str*, *applicable_date: str*, *data_subtype: str = None*, *additional_metadata: dict = None*)

Write metadata record to DynamoDB for a single file

libera_utils.db.dynamodb_utils.**create_ddb_metadata_file_item**(*filename: str*, *algorithm_version: str*, *include_archive_time: bool = False*, *additional_metadata: dict = None*)

Write metadata record to DynamoDB for a single file

libera_utils.db.dynamodb_utils.**get_dynamodb_table**(*dynamo_table_name: str*)

Get the DynamoDB table

3.1.6 libera_utils.io

Modules

caching	Module containing code to manage local file caching
filenaming	Module for file naming utilities
manifest	Module for manifest file handling
netcdf	Module containing utilities for writing Libera-conforming NetCDF4 data products
product_definition	Data Product configuration and writing for Libera NetCDF4 data product files
smart_open	Module for smart_open

libera_utils.io.caching

Module containing code to manage local file caching

Functions

empty_local_cache_dir()	Remove all cached files in the local cache.
get_local_cache_dir()	Determine where to cache files based on the system and installed package version.

libera_utils.io.caching.empty_local_cache_dir

libera_utils.io.caching.**empty_local_cache_dir**()

Remove all cached files in the local cache.

Returns
List of removed files

Return type
list

libera_utils.io.caching.get_local_cache_dir

libera_utils.io.caching.**get_local_cache_dir**()

Determine where to cache files based on the system and installed package version.

Returns
Path to the cache directory for this version of this package on the current system

Return type
pathlib.Path

libera_utils.io.caching.**empty_local_cache_dir**()

Remove all cached files in the local cache.

Returns
List of removed files

Return type
list

libera_utils.io.caching.**get_local_cache_dir**()

Determine where to cache files based on the system and installed package version.

Returns
Path to the cache directory for this version of this package on the current system

Return type
pathlib.Path

libera_utils.io.filenaming

Module for file naming utilities

Functions

format_semantic_version(semantic_version)	Formats a semantic version string X.Y.Z into a filename-compatible string like VX-Y-Z, for X = major version, Y = minor version, Z = patch.
get_current_version_str(package_name)	Retrieve the current version of a (algorithm) package and format it for inclusion in a filename

libera_utils.io.filenaming.format_semantic_version

libera_utils.io.filenaming.**format_semantic_version**(*semantic_version: str*) → str

Formats a semantic version string X.Y.Z into a filename-compatible string like VX-Y-Z, for X = major version, Y = minor version, Z = patch.

Result is uppercase. Release candidate suffixes are allowed as no strict checking is done on the contents of X, Y, or Z. e.g. 1.2.3rc1 becomes V1-2-3RC1

Parameters
semantic_version (*str*) – String matching X.Y.Z where X, Y and Z are integers of any length

Return type
str

libera_utils.io.filenaming.get_current_version_str

libera_utils.io.filenaming.**get_current_version_str**(*package_name: str*) → str

Retrieve the current version of a (algorithm) package and format it for inclusion in a filename

Parameters

package_name (*str*) – Package for which to retrieve a version string. This should be your algorithm package and it must use a semantic versioning scheme, configured in project metadata.

Returns

Version string in format vM1m2p3

Return type

str

Classes

AbstractDataProductFilename(*args, **kwargs)	Abstract base class for data product filenames.
AbstractValidFilename(*args, **kwargs)	Filename class that ensures validity of a filename based on regex pattern.
L0Filename(*args, **kwargs)	Filename validation class for L0 Production Data Set (PDS) files from EDOS.
LiberaDataProductFilename(*args, **kwargs)	Filename validation class for Libera SDC data products.
ManifestFilename(*args, **kwargs)	Class for naming manifest files

libera_utils.io.filenaming.AbstractDataProductFilename

class libera_utils.io.filenaming.**AbstractDataProductFilename**(**args*, ***kwargs*)

Bases: *AbstractValidFilename*

Abstract base class for data product filenames.

This class adds the data product specific requirements that all data products must have: a processing step ID and a data product ID. For example, an L0Filename or a LiberaDataProductFilename are both AbstractDataProductFilenames.

Attributes

archive_prefix
Property that contains the generated prefix used for archiving, when applicable

data_product_id
Property that contains the DataProductIdentifier for this file type

filename_parts
Property that contains a namespace of filename parts

path
Property containing the file path

Methods

from_file_path(*args, **kwargs)	Factory method to produce an AbstractValidFilename from a valid Libera file path (str or Path)

continues on next page

Table 55 – continued from previous page

from_filename_parts(*args, **kwargs)	Abstract method that must be implemented to provide hinting for required parts
generate_prefixed_path(parent_path)	Generates an absolute path of the form {parent_path}/{prefix_structure}/{file_basename} The parent_path can be an S3 bucket or an absolute local filepath (must start with /)
regex_match(path)	Parse and validate a given path against class-attribute defined regex

__init__(**args*, ***kwargs*)

Methods

from_file_path(*args, **kwargs)	Factory method to produce an AbstractValidFilename from a valid Libera file path (str or Path)
from_filename_parts(*args, **kwargs)	Abstract method that must be implemented to provide hinting for required parts
generate_prefixed_path(parent_path)	Generates an absolute path of the form {parent_path}/{prefix_structure}/{file_basename} The parent_path can be an S3 bucket or an absolute local filepath (must start with /)
regex_match(path)	Parse and validate a given path against class-attribute defined regex

Attributes

archive_prefix	Property that contains the generated prefix used for archiving, when applicable
data_product_id	Property that contains the DataProductIdentifier for this file type
filename_parts	Property that contains a namespace of filename parts
path	Property containing the file path

abstractmethod classmethod _format_filename_parts(**args: Any*, ***kwargs: Any*)

Format parts into a filename

Note: When this is implemented by concrete classes, *args and **kwargs become specific parameters

classmethod _from_filename_parts(*, *basepath: str | Path | S3Path | None = None*, ***parts: Any*)

Create instance from filename parts.

The part kwarg names are named according to the regex for the file type.

Parameters

- **basepath** (*Union[str, Path, S3Path], Optional*) – Allows prepending a basepath or prefix.
- **parts** (*Any*) – Passed directly to _format_filename_parts. This is a dict of variable kwargs that will differ in each filename class based on the required parts for that particular filename type.

Return type
AbstractValidFilename

abstractmethod _parse_filename_parts()

Parse the filename parts into objects from regex matched strings

Returns
namespace object containing filename parts as parsed objects

Return type
types.SimpleNamespace

abstract property archive_prefix: str

Property that contains the generated prefix used for archiving, when applicable

abstract property data_product_id: *DataProductIdentifier*

Property that contains the DataProductIdentifier for this file type

property filename_parts

Property that contains a namespace of filename parts

classmethod from_file_path(**args*, ***kwargs*)

Factory method to produce an AbstractValidFilename from a valid Libera file path (str or Path)

abstractmethod classmethod from_filename_parts(**args: Any*, ***kwargs: Any*)

Abstract method that must be implemented to provide hinting for required parts

generate_prefixed_path(*parent_path: str | CloudPath | Path*) → CloudPath | Path

Generates an absolute path of the form {parent_path}/{prefix_structure}/{file_basename} The parent_path can be an S3 bucket or an absolute local filepath (must start with /)

Parameters
parent_path (*Union[str, Path, S3Path]*) – Absolute path to the parent directory or S3 bucket prefix. The generated path prefix is appended to the parent path and followed by the file basename.

Return type
pathlib.Path or cloudpathlib.s3.s3path.S3Path

property path: CloudPath | Path

Property containing the file path

regex_match(*path: CloudPath | Path*)

Parse and validate a given path against class-attribute defined regex

Parameters
path (*Union[Path, CloudPath]*) – Path to validate

Returns
Match group dict of filename parts

Return type
dict

libera_utils.io.filenaming.AbstractValidFilename

class libera_utils.io.filenaming.**AbstractValidFilename**(**args*, ***kwargs*)

Bases: ABC

Filename class that ensures validity of a filename based on regex pattern.

Notes

- This is an abstract base class that must be inherited by concrete filename classes.
- This class internally stores a CloudPath or Path object in the *path* property (composition).

Attributes

archive_prefix
Property that contains the generated prefix used for archiving, when applicable

filename_parts
Property that contains a namespace of filename parts

path
Property containing the file path

Methods

from_file_path(*args, **kwargs)	Factory method to produce an AbstractValidFilename from a valid Libera file path (str or Path)
from_filename_parts(*args, **kwargs)	Abstract method that must be implemented to provide hinting for required parts
generate_prefixed_path(parent_path)	Generates an absolute path of the form {parent_path}/{prefix_structure}/{file_basename} The parent_path can be an S3 bucket or an absolute local filepath (must start with /)
regex_match(path)	Parse and validate a given path against class-attribute defined regex

__init__(**args*, ***kwargs*)

Methods

from_file_path(*args, **kwargs)	Factory method to produce an AbstractValidFilename from a valid Libera file path (str or Path)
from_filename_parts(*args, **kwargs)	Abstract method that must be implemented to provide hinting for required parts
generate_prefixed_path(parent_path)	Generates an absolute path of the form {parent_path}/{prefix_structure}/{file_basename} The parent_path can be an S3 bucket or an absolute local filepath (must start with /)
regex_match(path)	Parse and validate a given path against class-attribute defined regex

Attributes

archive_prefix	Property that contains the generated prefix used for archiving, when applicable
filename_parts	Property that contains a namespace of filename parts
path	Property containing the file path

abstractmethod classmethod _format_filename_parts(**args: Any*, ***kwargs: Any*)

Format parts into a filename

Note: When this is implemented by concrete classes, *args and **kwargs become specific parameters

classmethod _from_filename_parts(*, *basepath: str | Path | S3Path | None = None*, ***parts: Any*)

Create instance from filename parts.

The part kwarg names are named according to the regex for the file type.

Parameters

- **basepath** (*Union[str, Path, S3Path], Optional*) – Allows prepending a basepath or prefix.
- **parts** (*Any*) – Passed directly to _format_filename_parts. This is a dict of variable kwargs that will differ in each filename class based on the required parts for that particular filename type.

Return type
AbstractValidFilename

abstractmethod _parse_filename_parts()

Parse the filename parts into objects from regex matched strings

Returns
namespace object containing filename parts as parsed objects

Return type
types.SimpleNamespace

abstract property archive_prefix: str

Property that contains the generated prefix used for archiving, when applicable

property filename_parts

Property that contains a namespace of filename parts

classmethod from_file_path(**args*, ***kwargs*)

Factory method to produce an AbstractValidFilename from a valid Libera file path (str or Path)

abstractmethod classmethod from_filename_parts(**args: Any*, ***kwargs: Any*)

Abstract method that must be implemented to provide hinting for required parts

generate_prefixed_path(*parent_path: str | CloudPath | Path*) → CloudPath | Path

Generates an absolute path of the form {parent_path}/{prefix_structure}/{file_basename} The parent_path can be an S3 bucket or an absolute local filepath (must start with /)

Parameters
parent_path (*Union[str, Path, S3Path]*) – Absolute path to the parent directory or S3 bucket prefix. The generated path prefix is appended to the parent path and followed by the file basename.

Return type
pathlib.Path or cloudpathlib.s3.s3path.S3Path

property path: CloudPath | Path

Property containing the file path

regex_match(*path: CloudPath | Path*)

Parse and validate a given path against class-attribute defined regex

Parameters
path (*Union[Path, CloudPath]*) – Path to validate

Returns
Match group dict of filename parts

Return type
dict

libera_utils.io.filenaming.L0Filename

class libera_utils.io.filenaming.**L0Filename**(**args*, ***kwargs*)

Bases: *AbstractDataProductFilename*

Filename validation class for L0 Production Data Set (PDS) files from EDOS.

Attributes

archive_prefix
Property that contains the generated prefix for L0 archiving

data_product_id
Property that contains the DataProductIdentifier for this file type

filename_parts
Property that contains a namespace of filename parts

path
Property containing the file path

Methods

from_file_path(*args, **kwargs)	Factory method to produce an AbstractValidFilename from a valid Libera file path (str or Path)
from_filename_parts(*, id_char, scid, ...[, ...])	Create instance from filename parts
generate_prefixed_path(parent_path)	Generates an absolute path of the form {parent_path}/{prefix_structure}/{file_basename} The parent_path can be an S3 bucket or an absolute local filepath (must start with /)
regex_match(path)	Parse and validate a given path against class-attribute defined regex

__init__(**args*, ***kwargs*)

Methods

from_file_path(*args, **kwargs)	Factory method to produce an AbstractValidFilename from a valid Libera file path (str or Path)
from_filename_parts(*, id_char, scid, ...[, ...])	Create instance from filename parts
generate_prefixed_path(parent_path)	Generates an absolute path of the form {parent_path}/{prefix_structure}/{file_basename} The parent_path can be an S3 bucket or an absolute local filepath (must start with /)

continues on next page

Table 62 – continued from previous page

regex_match(path)	Parse and validate a given path against class-attribute defined regex

Attributes

archive_prefix	Property that contains the generated prefix for L0 archiving
data_product_id	Property that contains the DataProductIdentifier for this file type
filename_parts	Property that contains a namespace of filename parts
path	Property containing the file path

classmethod _format_filename_parts(*, *id_char: str*, *scid: int*, *first_apid: int*, *fill: str*, *created_time: datetime*, *numeric_id: int*, *file_number: int*, *extension: str*, *signal: str | None = None*)

Construct a path from filename parts

Parameters

- **id_char** (*str*) – Either P (for PDS files, Construction Records) or X (for Delivery Records)
- **scid** (*int*) – Spacecraft ID
- **first_apid** (*int*) – First APID in the file
- **fill** (*str*) – Custom string up to 14 characters long
- **created_time** (*datetime.datetime*) – Creation time of the file
- **numeric_id** (*int*) – Data set ID, 0-9, one digit
- **file_number** (*str*) – File number within the data set. Construction records are always file number zero.
- **extension** (*str*) – File name extension. Either PDR or PDS
- **signal** (*Optional[str], Optional*) – Optional signal suffix. Always '.XFR'

Returns

Formatted filename

Return type

str

classmethod _from_filename_parts(*, *basepath: str | Path | S3Path | None = None*, ***parts: Any*)

Create instance from filename parts.

The part kwarg names are named according to the regex for the file type.

Parameters

- **basepath** (*Union[str, Path, S3Path], Optional*) – Allows prepending a basepath or prefix.
- **parts** (*Any*) – Passed directly to _format_filename_parts. This is a dict of variable kwargs that will differ in each filename class based on the required parts for that particular filename type.

Return type
AbstractValidFilename

_parse_filename_parts()

Parse the filename parts into objects from regex matched strings

Returns
namespace object containing filename parts as parsed objects

Return type
types.SimpleNamespace

property archive_prefix: str

Property that contains the generated prefix for L0 archiving

property data_product_id: *DataProductIdentifier*

Property that contains the DataProductIdentifier for this file type

property filename_parts

Property that contains a namespace of filename parts

classmethod from_file_path(**args*, ***kwargs*)

Factory method to produce an AbstractValidFilename from a valid Libera file path (str or Path)

classmethod from_filename_parts(**, id_char: str, scid: int, first_apid: int, fill: str, created_time: datetime, numeric_id: int, file_number: int, extension: str, signal: str | None = None, basepath: str | Path | S3Path | None = None*)

Create instance from filename parts

This method exists primarily to expose typehinting to the user for use with the generic _from_filename_parts. The part names are named according to the regex for the file type.

Parameters

- **id_char** (*str*) – Either P (for PDS files, Construction Records) or X (for Delivery Records)
- **scid** (*int*) – Spacecraft ID
- **first_apid** (*int*) – First APID in the file
- **fill** (*str*) – Custom string up to 14 characters long
- **created_time** (*datetime.datetime*) – Creation time of the file
- **numeric_id** (*int*) – Data set ID, 0-9, one digit
- **file_number** (*str*) – File number within the data set. Construction records are always file number zero.
- **extension** (*str*) – File name extension. Either PDR or PDS
- **signal** (*Optional[str]*) – Optional signal suffix. Always '.XFR'
- **basepath** (*Optional[Union[str, Path, S3Path]]*) – Allows prepending a basepath or prefix.

Return type
L0Filename

generate_prefixed_path(*parent_path: str | CloudPath | Path*) → CloudPath | Path

Generates an absolute path of the form {parent_path}/{prefix_structure}/{file_basename} The parent_path can be an S3 bucket or an absolute local filepath (must start with /)

Parameters
parent_path (*Union[str, Path, S3Path]*) – Absolute path to the parent directory or S3 bucket prefix. The generated path prefix is appended to the parent path and followed by the file basename.

Return type
pathlib.Path or cloudpathlib.s3.s3path.S3Path

property path: CloudPath | Path

Property containing the file path

regex_match(*path: CloudPath | Path*)

Parse and validate a given path against class-attribute defined regex

Parameters
path (*Union[Path, CloudPath]*) – Path to validate

Returns
Match group dict of filename parts

Return type
dict

libera_utils.io.filenaming.LiberaDataProductFilename

class libera_utils.io.filenaming.**LiberaDataProductFilename**(**args*, ***kwargs*)

Bases: *AbstractDataProductFilename*

Filename validation class for Libera SDC data products.

Attributes

applicable_date
Property that returns the applicable date based on the midpoint of start and end times.

archive_prefix
Property that contains the generated prefix for L1B and L2 archiving

data_product_id
Property that contains the DataProductIdentifier for this file type

filename_parts
Property that contains a namespace of filename parts

path
Property containing the file path

processing_step_id
Property that contains the ProcessingStepIdentifier that generates this file

Methods

from_file_path(*args, **kwargs)	Factory method to produce an AbstractValidFilename from a valid Libera file path (str or Path)
from_filename_parts(*, product_name, ...[, ...])	Create instance from filename parts.
generate_prefixed_path(parent_path)	Generates an absolute path of the form {parent_path}/{prefix_structure}/{file_basename} The parent_path can be an S3 bucket or an absolute local filepath (must start with /)

continues on next page

Table 64 – continued from previous page

regex_match(path)	Parse and validate a given path against class-attribute defined regex

__init__(**args*, ***kwargs*)

Methods

from_file_path(*args, **kwargs)	Factory method to produce an AbstractValidFilename from a valid Libera file path (str or Path)
from_filename_parts(*, product_name, ...[, ...])	Create instance from filename parts.
generate_prefixed_path(parent_path)	Generates an absolute path of the form {parent_path}/{prefix_structure}/{file_basename} The parent_path can be an S3 bucket or an absolute local filepath (must start with /)
regex_match(path)	Parse and validate a given path against class-attribute defined regex

Attributes

applicable_date	Property that returns the applicable date based on the midpoint of start and end times.
archive_prefix	Property that contains the generated prefix for L1B and L2 archiving
data_product_id	Property that contains the DataProductIdentifier for this file type
filename_parts	Property that contains a namespace of filename parts
path	Property containing the file path
processing_step_id	Property that contains the ProcessingStepIdentifier that generates this file

classmethod _format_filename_parts(*, *data_level: str*, *product_name: str*, *version: str*, *utc_start: datetime*, *utc_end: datetime*, *revision: datetime*, *extension: str*)

Construct a path from filename parts

Parameters

- **data_level** (*str*) – L1B or L2
- **product_name** (*str*) – Libera instrument, cam or rad for L1B and cloud-fraction etc. for L2. May contain anything except for underscores.
- **version** (*str*) – Software version that the file was created with. Corresponds to the algorithm version as determined by the algorithm software.
- **utc_start** (*datetime.datetime*) – First timestamp in the SPK
- **utc_end** (*datetime.datetime*) – Last timestamp in the SPK
- **revision** (*datetime.datetime*) – Time when the file was created.
- **extension** (*str*) – File extension (.nc or .h5)

Returns

Formatted filename

Return type
str

classmethod _from_filename_parts(*, *basepath: str | Path | S3Path | None = None*, ***parts: Any*)

Create instance from filename parts.

The part kwarg names are named according to the regex for the file type.

Parameters

- **basepath** (*Union[str, Path, S3Path], Optional*) – Allows prepending a basepath or prefix.
- **parts** (*Any*) – Passed directly to _format_filename_parts. This is a dict of variable kwargs that will differ in each filename class based on the required parts for that particular filename type.

Return type
AbstractValidFilename

_parse_filename_parts()

Parse the filename parts into objects from regex matched strings

Returns
namespace object containing filename parts as parsed objects

Return type
types.SimpleNamespace

property applicable_date: date

Property that returns the applicable date based on the midpoint of start and end times.

Issues a warning if the time range covers more than 24 hours.

Returns
The date of the midpoint between utc_start and utc_end

Return type
datetime.date

property archive_prefix: str

Property that contains the generated prefix for L1B and L2 archiving

property data_product_id: *DataProductIdentifier*

Property that contains the DataProductIdentifier for this file type

property filename_parts

Property that contains a namespace of filename parts

classmethod from_file_path(**args*, ***kwargs*)

Factory method to produce an AbstractValidFilename from a valid Libera file path (str or Path)

classmethod from_filename_parts(*, *product_name: str | DataProductIdentifier*, *version: str*, *utc_start: datetime*, *utc_end: datetime*, *data_level: str | DataLevel | None = None*, *revision: datetime = datetime.datetime(2025, 11, 7, 17, 16, 50, 192919, tzinfo=datetime.timezone.utc)*, *extension: str | None = None*, *basepath: str | Path | S3Path | None = None*)

Create instance from filename parts. All keyword arguments other than basepath are required!

This method exists primarily to expose typehinting to the user for use with the generic _from_filename_parts. The part names are named according to the regex for the file type.

Parameters

- **data_level** (*str | DataLevel | None*) – L1B or L2 identifying the level of the data product. Default None will infer the data level from the product name (DataProductIdentifier)
- **product_name** (*str | DataProductIdentifier*) – Product type. e.g. CF-RAD for L2 or RAD-4CH for L1B. May contain anything except for underscores.
- **version** (*str*) – Software version that the file was created with. Corresponds to the algorithm version as determined by the algorithm software.
- **utc_start** (*datetime.datetime*) – First timestamp in the SPK
- **utc_end** (*datetime.datetime*) – Last timestamp in the SPK
- **revision** (*datetime.datetime*) – Time when the file was created. Default is now in UTC time.
- **extension** (*str | None*) – File extension. Default None will infer extension based on product_name.
- **basepath** (*Optional[Union[str, Path, S3Path]]*) – Allows prepending a basepath or prefix.

Return type
LiberaDataProductFilename

generate_prefixed_path(*parent_path: str | CloudPath | Path*) → CloudPath | Path

Generates an absolute path of the form {parent_path}/{prefix_structure}/{file_basename} The parent_path can be an S3 bucket or an absolute local filepath (must start with /)

Parameters
parent_path (*Union[str, Path, S3Path]*) – Absolute path to the parent directory or S3 bucket prefix. The generated path prefix is appended to the parent path and followed by the file basename.

Return type
pathlib.Path or cloudpathlib.s3.s3path.S3Path

property path: CloudPath | Path

Property containing the file path

property processing_step_id: *ProcessingStepIdentifier* | None

Property that contains the ProcessingStepIdentifier that generates this file

regex_match(*path: CloudPath | Path*)

Parse and validate a given path against class-attribute defined regex

Parameters
path (*Union[Path, CloudPath]*) – Path to validate

Returns
Match group dict of filename parts

Return type
dict

libera_utils.io.filenaming.ManifestFilename

class libera_utils.io.filenaming.**ManifestFilename**(**args*, ***kwargs*)

Bases: *AbstractValidFilename*

Class for naming manifest files

Attributes

archive_prefix
Manifests are not archived like data products, but for convenience and ease of debugging they will be kept in the dropbox bucket by input/output and day they were made.

filename_parts
Property that contains a namespace of filename parts

path
Property containing the file path

Methods

from_file_path(*args, **kwargs)	Factory method to produce an AbstractValidFilename from a valid Libera file path (str or Path)
from_filename_parts(manifest_type, ulid_code)	Create instance from filename parts.
generate_prefixed_path(parent_path)	Generates an absolute path of the form {parent_path}/{prefix_structure}/{file_basename} The parent_path can be an S3 bucket or an absolute local filepath (must start with /)
regex_match(path)	Parse and validate a given path against class-attribute defined regex

__init__(**args*, ***kwargs*)

Methods

from_file_path(*args, **kwargs)	Factory method to produce an AbstractValidFilename from a valid Libera file path (str or Path)
from_filename_parts(manifest_type, ulid_code)	Create instance from filename parts.
generate_prefixed_path(parent_path)	Generates an absolute path of the form {parent_path}/{prefix_structure}/{file_basename} The parent_path can be an S3 bucket or an absolute local filepath (must start with /)
regex_match(path)	Parse and validate a given path against class-attribute defined regex

Attributes

archive_prefix	Manifests are not archived like data products, but for convenience and ease of debugging they will be kept in the dropbox bucket by input/output and day they were made.
filename_parts	Property that contains a namespace of filename parts

continues on next page

Table 69 – continued from previous page

path	Property containing the file path

classmethod _format_filename_parts(*manifest_type: ManifestType*, *ulid_code: ULID*)

Construct a path from filename parts

Parameters

- **manifest_type** (*ManifestType*) – Input or output
- **ulid_code** (*ulid.ULID*) – ULID code for use in filename parts

Returns
Formatted filename

Return type
str

classmethod _from_filename_parts(*, *basepath: str | Path | S3Path | None = None*, ***parts: Any*)

Create instance from filename parts.

The part kwarg names are named according to the regex for the file type.

Parameters

- **basepath** (*Union[str, Path, S3Path], Optional*) – Allows prepending a basepath or prefix.
- **parts** (*Any*) – Passed directly to _format_filename_parts. This is a dict of variable kwargs that will differ in each filename class based on the required parts for that particular filename type.

Return type
AbstractValidFilename

_parse_filename_parts()

Parse the filename parts into objects from regex matched strings

Returns
namespace object containing filename parts as parsed objects

Return type
types.SimpleNamespace

property archive_prefix: str

Manifests are not archived like data products, but for convenience and ease of debugging they will be kept in the dropbox bucket by input/output and day they were made. This is used by the step function clean up function in the CDK. # Generate prefix structure # <manifest_type>/<year>/<month>/<day>

property filename_parts

Property that contains a namespace of filename parts

classmethod from_file_path(**args*, ***kwargs*)

Factory method to produce an AbstractValidFilename from a valid Libera file path (str or Path)

classmethod from_filename_parts(*manifest_type: ManifestType*, *ulid_code: ULID*, *basepath: str | Path | S3Path | None = None*)

Create instance from filename parts.

This method exists primarily to expose typehinting to the user for use with the generic _from_filename_parts. The part names are named according to the regex for the file type.

Parameters

- **manifest_type** (ManifestType) – Input or output
- **ulid_code** (*ulid.ULID*) – ULID code for use in filename parts
- **basepath** (*Optional[Union[str, Path, S3Path]]*) – Allows prepending a basepath or prefix.

Return type
ManifestFilename

generate_prefixed_path(*parent_path: str | CloudPath | Path*) → CloudPath | Path

Generates an absolute path of the form {parent_path}/{prefix_structure}/{file_basename} The parent_path can be an S3 bucket or an absolute local filepath (must start with /)

Parameters
parent_path (*Union[str, Path, S3Path]*) – Absolute path to the parent directory or S3 bucket prefix. The generated path prefix is appended to the parent path and followed by the file basename.

Return type
pathlib.Path or cloudpathlib.s3.s3path.S3Path

property path: CloudPath | Path

Property containing the file path

regex_match(*path: CloudPath | Path*)

Parse and validate a given path against class-attribute defined regex

Parameters
path (*Union[Path, CloudPath]*) – Path to validate

Returns
Match group dict of filename parts

Return type
dict

class libera_utils.io.filenaming.**AbstractDataProductFilename**(**args*, ***kwargs*)

Abstract base class for data product filenames.

This class adds the data product specific requirements that all data products must have: a processing step ID and a data product ID. For example, an L0Filename or a LiberaDataProductFilename are both AbstractDataProductFilenames.

Attributes

archive_prefix
Property that contains the generated prefix used for archiving, when applicable

data_product_id
Property that contains the DataProductIdentifier for this file type

filename_parts
Property that contains a namespace of filename parts

path
Property containing the file path

Methods

from_file_path(*args, **kwargs)	Factory method to produce an AbstractValidFilename from a valid Libera file path (str or Path)
from_filename_parts(*args, **kwargs)	Abstract method that must be implemented to provide hinting for required parts
generate_prefixed_path(parent_path)	Generates an absolute path of the form {parent_path}/{prefix_structure}/{file_basename} The parent_path can be an S3 bucket or an absolute local filepath (must start with /)
regex_match(path)	Parse and validate a given path against class-attribute defined regex

abstract property data_product_id: *DataProductIdentifier*

Property that contains the DataProductIdentifier for this file type

class libera_utils.io.filenaming.**AbstractValidFilename**(**args*, ***kwargs*)

Filename class that ensures validity of a filename based on regex pattern.

Notes

- This is an abstract base class that must be inherited by concrete filename classes.
- This class internally stores a CloudPath or Path object in the *path* property (composition).

Attributes

archive_prefix
Property that contains the generated prefix used for archiving, when applicable

filename_parts
Property that contains a namespace of filename parts

path
Property containing the file path

Methods

from_file_path(*args, **kwargs)	Factory method to produce an AbstractValidFilename from a valid Libera file path (str or Path)
from_filename_parts(*args, **kwargs)	Abstract method that must be implemented to provide hinting for required parts
generate_prefixed_path(parent_path)	Generates an absolute path of the form {parent_path}/{prefix_structure}/{file_basename} The parent_path can be an S3 bucket or an absolute local filepath (must start with /)
regex_match(path)	Parse and validate a given path against class-attribute defined regex

abstractmethod classmethod _format_filename_parts(**args: Any*, ***kwargs: Any*)

Format parts into a filename

Note: When this is implemented by concrete classes, *args and **kwargs become specific parameters

classmethod _from_filename_parts(*, *basepath: str | Path | S3Path | None = None*, ***parts: Any*)

Create instance from filename parts.

The part kwarg names are named according to the regex for the file type.

Parameters

- **basepath** (*Union[str, Path, S3Path], Optional*) – Allows prepending a basepath or prefix.
- **parts** (*Any*) – Passed directly to _format_filename_parts. This is a dict of variable kwargs that will differ in each filename class based on the required parts for that particular filename type.

Return type
AbstractValidFilename

abstractmethod _parse_filename_parts()

Parse the filename parts into objects from regex matched strings

Returns
namespace object containing filename parts as parsed objects

Return type
types.SimpleNamespace

abstract property archive_prefix: str

Property that contains the generated prefix used for archiving, when applicable

property filename_parts

Property that contains a namespace of filename parts

classmethod from_file_path(**args*, ***kwargs*)

Factory method to produce an AbstractValidFilename from a valid Libera file path (str or Path)

abstractmethod classmethod from_filename_parts(**args: Any*, ***kwargs: Any*)

Abstract method that must be implemented to provide hinting for required parts

generate_prefixed_path(*parent_path: str | CloudPath | Path*) → CloudPath | Path

Generates an absolute path of the form {parent_path}/{prefix_structure}/{file_basename} The parent_path can be an S3 bucket or an absolute local filepath (must start with /)

Parameters
parent_path (*Union[str, Path, S3Path]*) – Absolute path to the parent directory or S3 bucket prefix. The generated path prefix is appended to the parent path and followed by the file basename.

Return type
pathlib.Path or cloudpathlib.s3.s3path.S3Path

property path: CloudPath | Path

Property containing the file path

regex_match(*path: CloudPath | Path*)

Parse and validate a given path against class-attribute defined regex

Parameters
path (*Union[Path, CloudPath]*) – Path to validate

Returns
Match group dict of filename parts

Return type
dict

class libera_utils.io.filenaming.**L0Filename**(**args*, ***kwargs*)

Filename validation class for L0 Production Data Set (PDS) files from EDOS.

Attributes

archive_prefix
Property that contains the generated prefix for L0 archiving

data_product_id
Property that contains the DataProductIdentifier for this file type

filename_parts
Property that contains a namespace of filename parts

path
Property containing the file path

Methods

from_file_path(*args, **kwargs)	Factory method to produce an AbstractValidFilename from a valid Libera file path (str or Path)
from_filename_parts(*, id_char, scid, ...[, ...])	Create instance from filename parts
generate_prefixed_path(parent_path)	Generates an absolute path of the form {parent_path}/{prefix_structure}/{file_basename} The parent_path can be an S3 bucket or an absolute local filepath (must start with /)
regex_match(path)	Parse and validate a given path against class-attribute defined regex

classmethod **_format_filename_parts**(*, *id_char: str*, *scid: int*, *first_apid: int*, *fill: str*, *created_time: datetime*, *numeric_id: int*, *file_number: int*, *extension: str*, *signal: str | None = None*)

Construct a path from filename parts

Parameters

- **id_char** (*str*) – Either P (for PDS files, Construction Records) or X (for Delivery Records)
- **scid** (*int*) – Spacecraft ID
- **first_apid** (*int*) – First APID in the file
- **fill** (*str*) – Custom string up to 14 characters long
- **created_time** (*datetime.datetime*) – Creation time of the file
- **numeric_id** (*int*) – Data set ID, 0-9, one digit
- **file_number** (*str*) – File number within the data set. Construction records are always file number zero.
- **extension** (*str*) – File name extension. Either PDR or PDS
- **signal** (*Optional[str], Optional*) – Optional signal suffix. Always '.XFR'

Returns
Formatted filename

Return type
str

_parse_filename_parts()

Parse the filename parts into objects from regex matched strings

Returns
namespace object containing filename parts as parsed objects

Return type
types.SimpleNamespace

property archive_prefix: str

Property that contains the generated prefix for L0 archiving

property data_product_id: *DataProductIdentifier*

Property that contains the DataProductIdentifier for this file type

classmethod from_filename_parts(*, *id_char: str*, *scid: int*, *first_apid: int*, *fill: str*, *created_time: datetime*, *numeric_id: int*, *file_number: int*, *extension: str*, *signal: str | None = None*, *basepath: str | Path | S3Path | None = None*)

Create instance from filename parts

This method exists primarily to expose typehinting to the user for use with the generic _from_filename_parts. The part names are named according to the regex for the file type.

Parameters

- **id_char** (*str*) – Either P (for PDS files, Construction Records) or X (for Delivery Records)
- **scid** (*int*) – Spacecraft ID
- **first_apid** (*int*) – First APID in the file
- **fill** (*str*) – Custom string up to 14 characters long
- **created_time** (*datetime.datetime*) – Creation time of the file
- **numeric_id** (*int*) – Data set ID, 0-9, one digit
- **file_number** (*str*) – File number within the data set. Construction records are always file number zero.
- **extension** (*str*) – File name extension. Either PDR or PDS
- **signal** (*Optional[str]*) – Optional signal suffix. Always '.XFR'
- **basepath** (*Optional[Union[str, Path, S3Path]]*) – Allows prepending a basepath or prefix.

Return type
L0Filename

class libera_utils.io.filenaming.**LiberaDataProductFilename**(**args*, ***kwargs*)

Filename validation class for Libera SDC data products.

Attributes

applicable_date
Property that returns the applicable date based on the midpoint of start and end times.

archive_prefix
Property that contains the generated prefix for L1B and L2 archiving

data_product_id
Property that contains the DataProductIdentifier for this file type

filename_parts
Property that contains a namespace of filename parts

path
Property containing the file path

processing_step_id
Property that contains the ProcessingStepIdentifier that generates this file

Methods

from_file_path(*args, **kwargs)	Factory method to produce an AbstractValidFilename from a valid Libera file path (str or Path)
from_filename_parts(*, product_name, ...[, ...])	Create instance from filename parts.
generate_prefixed_path(parent_path)	Generates an absolute path of the form {parent_path}/{prefix_structure}/{file_basename} The parent_path can be an S3 bucket or an absolute local filepath (must start with /)
regex_match(path)	Parse and validate a given path against class-attribute defined regex

classmethod _format_filename_parts(*, *data_level: str*, *product_name: str*, *version: str*, *utc_start: datetime*, *utc_end: datetime*, *revision: datetime*, *extension: str*)

Construct a path from filename parts

Parameters

- **data_level** (*str*) – L1B or L2
- **product_name** (*str*) – Libera instrument, cam or rad for L1B and cloud-fraction etc. for L2. May contain anything except for underscores.
- **version** (*str*) – Software version that the file was created with. Corresponds to the algorithm version as determined by the algorithm software.
- **utc_start** (*datetime.datetime*) – First timestamp in the SPK
- **utc_end** (*datetime.datetime*) – Last timestamp in the SPK
- **revision** (*datetime.datetime*) – Time when the file was created.
- **extension** (*str*) – File extension (.nc or .h5)

Returns
Formatted filename

Return type
str

_parse_filename_parts()

Parse the filename parts into objects from regex matched strings

Returns
namespace object containing filename parts as parsed objects

Return type
types.SimpleNamespace

property applicable_date: *date*

Property that returns the applicable date based on the midpoint of start and end times.

Issues a warning if the time range covers more than 24 hours.

Returns
The date of the midpoint between utc_start and utc_end

Return type
datetime.date

property archive_prefix: *str*

Property that contains the generated prefix for L1B and L2 archiving

property data_product_id: *DataProductIdentifier*

Property that contains the DataProductIdentifier for this file type

classmethod from_filename_parts(*, *product_name: str | DataProductIdentifier*, *version: str*, *utc_start: datetime*, *utc_end: datetime*, *data_level: str | DataLevel | None = None*, *revision: datetime = datetime.datetime(2025, 11, 7, 17, 16, 50, 192919, tzinfo=datetime.timezone.utc)*, *extension: str | None = None*, *basepath: str | Path | S3Path | None = None*)

Create instance from filename parts. All keyword arguments other than basepath are required!

This method exists primarily to expose typehinting to the user for use with the generic _from_filename_parts. The part names are named according to the regex for the file type.

Parameters

- **data_level** (*str | DataLevel | None*) – L1B or L2 identifying the level of the data product. Default None will infer the data level from the product name (DataProductIdentifier)
- **product_name** (*str | DataProductIdentifier*) – Product type. e.g. CF-RAD for L2 or RAD-4CH for L1B. May contain anything except for underscores.
- **version** (*str*) – Software version that the file was created with. Corresponds to the algorithm version as determined by the algorithm software.
- **utc_start** (*datetime.datetime*) – First timestamp in the SPK
- **utc_end** (*datetime.datetime*) – Last timestamp in the SPK
- **revision** (*datetime.datetime*) – Time when the file was created. Default is now in UTC time.
- **extension** (*str | None*) – File extension. Default None will infer extension based on product_name.
- **basepath** (*Optional[Union[str, Path, S3Path]]*) – Allows prepending a basepath or prefix.

Return type
LiberaDataProductFilename

property processing_step_id: *ProcessingStepIdentifier* | *None*

Property that contains the ProcessingStepIdentifier that generates this file

class libera_utils.io.filenaming.**ManifestFilename**(**args*, ***kwargs*)

Class for naming manifest files

Attributes

archive_prefix
: Manifests are not archived like data products, but for convenience and ease of debugging they will be kept in the dropbox bucket by input/output and day they were made.

filename_parts
: Property that contains a namespace of filename parts

path
: Property containing the file path

Methods

from_file_path(*args, **kwargs)	Factory method to produce an AbstractValidFilename from a valid Libera file path (str or Path)
from_filename_parts(manifest_type, ulid_code)	Create instance from filename parts.
generate_prefixed_path(parent_path)	Generates an absolute path of the form {parent_path}/{prefix_structure}/{file_basename} The parent_path can be an S3 bucket or an absolute local filepath (must start with /)
regex_match(path)	Parse and validate a given path against class-attribute defined regex

classmethod _format_filename_parts(*manifest_type: ManifestType*, *ulid_code: ULID*)

Construct a path from filename parts

Parameters

- **manifest_type** (ManifestType) – Input or output
- **ulid_code** (*ulid.ULID*) – ULID code for use in filename parts

Returns
Formatted filename

Return type
str

_parse_filename_parts()

Parse the filename parts into objects from regex matched strings

Returns
namespace object containing filename parts as parsed objects

Return type
types.SimpleNamespace

property archive_prefix: str

Manifests are not archived like data products, but for convenience and ease of debugging they will be kept in the dropbox bucket by input/output and day they were made. This is used by the step function clean up function in the CDK. # Generate prefix structure # <manifest_type>/<year>/<month>/<day>

classmethod from_filename_parts(*manifest_type: ManifestType*, *ulid_code: ULID*, *basepath: str | Path | S3Path | None = None*)

Create instance from filename parts.

This method exists primarily to expose typehinting to the user for use with the generic _from_filename_parts. The part names are named according to the regex for the file type.

Parameters

- **manifest_type** (ManifestType) – Input or output
- **ulid_code** (*ulid.ULID*) – ULID code for use in filename parts
- **basepath** (*Optional[Union[str, Path, S3Path]]*) – Allows prepending a basepath or prefix.

Return type
ManifestFilename

libera_utils.io.filenaming.**_ensure_utc_timezone**(*dt_obj: datetime*) → datetime

Ensure datetime object has UTC timezone info.

If the datetime is timezone-naive, assume it is in UTC and add timezone info. If the datetime is timezone-aware, convert it to UTC.

Parameters
dt_obj (*datetime*) – Input datetime object

Returns
Timezone-aware datetime in UTC

Return type
datetime

libera_utils.io.filenaming.**format_semantic_version**(*semantic_version: str*) → str

Formats a semantic version string X.Y.Z into a filename-compatible string like VX-Y-Z, for X = major version, Y = minor version, Z = patch.

Result is uppercase. Release candidate suffixes are allowed as no strict checking is done on the contents of X, Y, or Z. e.g. 1.2.3rc1 becomes V1-2-3RC1

Parameters
semantic_version (*str*) – String matching X.Y.Z where X, Y and Z are integers of any length

Return type
str

libera_utils.io.filenaming.**get_current_version_str**(*package_name: str*) → str

Retrieve the current version of a (algorithm) package and format it for inclusion in a filename

Parameters
package_name (*str*) – Package for which to retrieve a version string. This should be your algorithm package and it must use a semantic versioning scheme, configured in project metadata.

Returns
Version string in format vM1m2p3

Return type
str

libera_utils.io.manifest

Module for manifest file handling

Functions

calculate_checksum(file)	Compute the checksum of the given file.
get_ulid_code(filename)	Get ULID code from filename.

libera_utils.io.manifest.calculate_checksum

libera_utils.io.manifest.**calculate_checksum**(*file: str | Path | S3Path*) → str

Compute the checksum of the given file.

libera_utils.io.manifest.get_ulid_code

libera_utils.io.manifest.**get_ulid_code**(*filename: str | Path | S3Path | ManifestFilename | None*) → ULID | None

Get ULID code from filename.

Classes

Manifest(*, manifest_type, files, ...)	Pydantic model for a manifest file.
ManifestFileRecord(*, filename, checksum)	Pydantic model for an individual data product file recorded within a manifest file.

libera_utils.io.manifest.Manifest

class libera_utils.io.manifest.**Manifest**(*, *manifest_type: ~libera_utils.constants.ManifestType*, *files: list[~libera_utils.io.manifest.ManifestFileRecord] = <factory>*, *configuration: dict[str, ~typing.Any] = <factory>*, *filename: ~libera_utils.io.filenaming.ManifestFilename | None = None*, *ulid_code: ~ulid.ULID | None = <factory>*)

Bases: BaseModel

Pydantic model for a manifest file.

Attributes

model_extra
Get extra fields set during validation.

model_fields_set
Returns the set of fields that have been explicitly set on this model instance.

Methods

add_desired_time_range(start_datetime, ...)	Add a time range to the configuration section of the manifest.
add_files(*files)	Add files to the manifest from filename
check_file_structure(file_structure, ...)	Check file structure, returning True if it is good.
copy(*[, include, exclude, update, deep])	Returns a copy of the model.
from_file(filepath)	Read a manifest file and return a Manifest object (factory method).
model_construct([_fields_set])	Creates a new instance of the *Model* class with validated data.
model_copy(*[, update, deep])	!!! abstract "Usage Documentation"
model_dump(*[, mode, include, exclude, ...])	!!! abstract "Usage Documentation"
model_dump_json(*[, indent, ensure_ascii, ...])	!!! abstract "Usage Documentation"
model_json_schema([by_alias, ref_template, ...])	Generates a JSON schema for a model class.

continues on next page

Table 77 – continued from previous page

model_parametrized_name(params)	Compute the class name for parametrizations of generic classes.
model_post_init(context, /)	Override this method to perform additional initialization after *__init__* and *model_construct*.
model_rebuild(*[, force, raise_errors, ...])	Try to rebuild the pydantic-core schema for the model.
model_validate(obj, *[, strict, extra, ...])	Validate a pydantic model instance.
model_validate_json(json_data, *[, strict, ...])	!!! abstract "Usage Documentation"
model_validate_strings(obj, *[, strict, ...])	Validate the given object with string data against the Pydantic model.
output_manifest_from_input_manifest(...)	Create Output manifest from input manifest file path, adds input files to output manifest configuration
serialize_filename(filename, _info)	Custom serializer for the manifest filename.
transform_filename(raw_filename)	Convert raw filename to ManifestFilename class if necessary.
transform_files(raw_list)	Allow for the incoming files list to have varying types.
validate_checksums()	Validate checksums of listed files
write(out_path[, filename])	Write a manifest file from a Manifest object (self).

construct
dict
from_orm
json
parse_file
parse_obj
parse_raw
schema
schema_json
update_forward_refs
validate

__init__(**data: *Any*) → None

Create a new model by parsing and validating input data from keyword arguments.

Raises [*ValidationError*][pydantic_core.ValidationError] if the input data cannot be validated to form a valid model.

self is explicitly positional-only to allow *self* as a field name.

Methods

add_desired_time_range(start_datetime, ...)	Add a time range to the configuration section of the manifest.
add_files(*files)	Add files to the manifest from filename
check_file_structure(file_structure, ...)	Check file structure, returning True if it is good.
from_file(filepath)	Read a manifest file and return a Manifest object (factory method).
output_manifest_from_input_manifest(...)	Create Output manifest from input manifest file path, adds input files to output manifest configuration
serialize_filename(filename, _info)	Custom serializer for the manifest filename.

continues on next page

Table 78 – continued from previous page

transform_filename(raw_filename)	Convert raw filename to ManifestFilename class if necessary.
transform_files(raw_list)	Allow for the incoming files list to have varying types.
validate_checksums()	Validate checksums of listed files
write(out_path[, filename])	Write a manifest file from a Manifest object (self).

Attributes

model_computed_fields	
model_config	Configuration for the model, should be a dictionary conforming to [*ConfigDict*][pydantic.config.ConfigDict].
model_extra	Get extra fields set during validation.
model_fields	
model_fields_set	Returns the set of fields that have been explicitly set on this model instance.
manifest_type	
files	
configuration	
filename	
ulid_code	

_generate_filename() → *ManifestFilename*

Generate a valid manifest filename

add_desired_time_range(*start_datetime: datetime*, *end_datetime: datetime*)

Add a time range to the configuration section of the manifest.

Parameters

- **start_datetime** (*datetime.datetime*) – The desired start time for the range of data in this manifest
- **end_datetime** (*datetime.datetime*) – The desired end time for the range of data in this manifest

Return type
None

add_files(**files: str | Path | S3Path*)

Add files to the manifest from filename

Parameters
files (*Union[str, Path, S3Path]*) – Path to the file to add to the manifest.

Return type
None

classmethod check_file_structure(*file_structure:* ManifestFileRecord, *existing_names: set[str]*, *existing_checksums: set[str]*) → bool

Check file structure, returning True if it is good.

classmethod from_file(*filepath: str | Path | S3Path*)

Read a manifest file and return a Manifest object (factory method).

Parameters
filepath (*Union[str, Path, S3Path]*) – Location of manifest file to read.

Returns
Pydantic model built from the json of the given manifest file.

Return type
Manifest

model_config: ClassVar[ConfigDict] = {'arbitrary_types_allowed': True}

Configuration for the model, should be a dictionary conforming to [*ConfigDict*][pydantic.config.ConfigDict].

property model_extra: dict[str, Any] | None

Get extra fields set during validation.

Returns
A dictionary of extra fields, or *None* if *config.extra* is not set to *"allow"*.

property model_fields_set: set[str]

Returns the set of fields that have been explicitly set on this model instance.

Returns

A set of strings representing the fields that have been set,
i.e. that were not filled from defaults.

classmethod output_manifest_from_input_manifest(*input_manifest: Path | S3Path | Manifest*) → *Manifest*

Create Output manifest from input manifest file path, adds input files to output manifest configuration

Parameters
input_manifest (*Union[Path, S3Path, 'Manifest']*) – An S3 or regular path to an input_manifest object, or the input manifest object itself

Returns
output_manifest – The newly created output manifest

Return type
Manifest

serialize_filename(*filename: str | Path | S3Path | ManifestFilename | None*, *_info*) → str

Custom serializer for the manifest filename.

classmethod transform_filename(*raw_filename: str | ManifestFilename | None*) → *ManifestFilename* | None

Convert raw filename to ManifestFilename class if necessary.

classmethod transform_files(*raw_list: list[dict | str | Path | S3Path | ManifestFileRecord] | None*) → list[*ManifestFileRecord*]

Allow for the incoming files list to have varying types. Convert to a standardized list of ManifestFileStructure.

validate_checksums() → None

Validate checksums of listed files

write(*out_path: str | Path | S3Path*, *filename: str = None*) → Path | S3Path

Write a manifest file from a Manifest object (self).

Parameters

- **out_path** (*Union[str, Path, S3Path]*) – Directory path to write to (directory being used loosely to refer also to an S3 bucket path).
- **filename** (*str, Optional*) – must be a valid manifest filename. If not provided, the method uses the objects internal filename attribute. If that is not set, then a filename is automatically generated.

Returns

The path where the manifest file is written.

Return type

Union[Path, S3Path]

libera_utils.io.manifest.ManifestFileRecord

class libera_utils.io.manifest.**ManifestFileRecord**(*, *filename: str*, *checksum: str*)

Bases: BaseModel

Pydantic model for an individual data product file recorded within a manifest file.

Attributes

model_extra

Get extra fields set during validation.

model_fields_set

Returns the set of fields that have been explicitly set on this model instance.

Methods

copy(*[, include, exclude, update, deep])	Returns a copy of the model.
model_construct([_fields_set])	Creates a new instance of the *Model* class with validated data.
model_copy(*[, update, deep])	!!! abstract "Usage Documentation"
model_dump(*[, mode, include, exclude, ...])	!!! abstract "Usage Documentation"
model_dump_json(*[, indent, ensure_ascii, ...])	!!! abstract "Usage Documentation"
model_json_schema([by_alias, ref_template, ...])	Generates a JSON schema for a model class.
model_parametrized_name(params)	Compute the class name for parametrizations of generic classes.
model_post_init(context, /)	Override this method to perform additional initialization after *__init__* and *model_construct*.
model_rebuild(*[, force, raise_errors, ...])	Try to rebuild the pydantic-core schema for the model.
model_validate(obj, *[, strict, extra, ...])	Validate a pydantic model instance.
model_validate_json(json_data, *[, strict, ...])	!!! abstract "Usage Documentation"
model_validate_strings(obj, *[, strict, ...])	Validate the given object with string data against the Pydantic model.

construct
dict
from_orm
json
parse_file
parse_obj
parse_raw
schema
schema_json
update_forward_refs
validate

__init__(**data: *Any*) → None

Create a new model by parsing and validating input data from keyword arguments.

Raises [*ValidationError*][pydantic_core.ValidationError] if the input data cannot be validated to form a valid model.

self is explicitly positional-only to allow *self* as a field name.

Attributes

model_computed_fields	
model_config	Configuration for the model, should be a dictionary conforming to [*ConfigDict*][pydantic.config.ConfigDict].
model_extra	Get extra fields set during validation.
model_fields	
model_fields_set	Returns the set of fields that have been explicitly set on this model instance.
filename	
checksum	

model_config: ClassVar[ConfigDict] = {}

Configuration for the model, should be a dictionary conforming to [*ConfigDict*][pydantic.config.ConfigDict].

property model_extra: dict[str, Any] | None

Get extra fields set during validation.

Returns

A dictionary of extra fields, or *None* if *config.extra* is not set to *"allow"*.

property model_fields_set: set[str]

Returns the set of fields that have been explicitly set on this model instance.

Returns

A set of strings representing the fields that have been set,
i.e. that were not filled from defaults.

Exceptions

ManifestError	Generic exception related to manifest file handling

libera_utils.io.manifest.ManifestError

exception libera_utils.io.manifest.**ManifestError**

Generic exception related to manifest file handling

class libera_utils.io.manifest.**Manifest**(*, *manifest_type: ~libera_utils.constants.ManifestType*, *files: list[~libera_utils.io.manifest.ManifestFileRecord] = <factory>*, *configuration: dict[str, ~typing.Any] = <factory>*, *filename: ~libera_utils.io.filenaming.ManifestFilename | None = None*, *ulid_code: ~ulid.ULID | None = <factory>*)

Pydantic model for a manifest file.

Attributes

model_extra
Get extra fields set during validation.

model_fields_set
Returns the set of fields that have been explicitly set on this model instance.

Methods

add_desired_time_range(start_datetime, ...)	Add a time range to the configuration section of the manifest.
add_files(*files)	Add files to the manifest from filename
check_file_structure(file_structure, ...)	Check file structure, returning True if it is good.
copy(*[, include, exclude, update, deep])	Returns a copy of the model.
from_file(filepath)	Read a manifest file and return a Manifest object (factory method).
model_construct([_fields_set])	Creates a new instance of the *Model* class with validated data.
model_copy(*[, update, deep])	!!! abstract "Usage Documentation"
model_dump(*[, mode, include, exclude, ...])	!!! abstract "Usage Documentation"
model_dump_json(*[, indent, ensure_ascii, ...])	!!! abstract "Usage Documentation"
model_json_schema([by_alias, ref_template, ...])	Generates a JSON schema for a model class.
model_parametrized_name(params)	Compute the class name for parametrizations of generic classes.
model_post_init(context, /)	Override this method to perform additional initialization after *__init__* and *model_construct*.
model_rebuild(*[, force, raise_errors, ...])	Try to rebuild the pydantic-core schema for the model.
model_validate(obj, *[, strict, extra, ...])	Validate a pydantic model instance.
model_validate_json(json_data, *[, strict, ...])	!!! abstract "Usage Documentation"
model_validate_strings(obj, *[, strict, ...])	Validate the given object with string data against the Pydantic model.

continues on next page

Table 83 – continued from previous page

output_manifest_from_input_manifest(...)	Create Output manifest from input manifest file path, adds input files to output manifest configuration
serialize_filename(filename, _info)	Custom serializer for the manifest filename.
transform_filename(raw_filename)	Convert raw filename to ManifestFilename class if necessary.
transform_files(raw_list)	Allow for the incoming files list to have varying types.
validate_checksums()	Validate checksums of listed files
write(out_path[, filename])	Write a manifest file from a Manifest object (self).

construct
dict
from_orm
json
parse_file
parse_obj
parse_raw
schema
schema_json
update_forward_refs
validate

_generate_filename() → *ManifestFilename*

Generate a valid manifest filename

add_desired_time_range(*start_datetime: datetime*, *end_datetime: datetime*)

Add a time range to the configuration section of the manifest.

Parameters

- **start_datetime** (*datetime.datetime*) – The desired start time for the range of data in this manifest
- **end_datetime** (*datetime.datetime*) – The desired end time for the range of data in this manifest

Return type
None

add_files(**files: str | Path | S3Path*)

Add files to the manifest from filename

Parameters
files (*Union[str, Path, S3Path]*) – Path to the file to add to the manifest.

Return type
None

classmethod check_file_structure(*file_structure:* ManifestFileRecord, *existing_names: set[str]*, *existing_checksums: set[str]*) → bool

Check file structure, returning True if it is good.

classmethod from_file(*filepath: str | Path | S3Path*)

Read a manifest file and return a Manifest object (factory method).

Parameters
filepath (*Union[str, Path, S3Path]*) – Location of manifest file to read.

Returns
Pydantic model built from the json of the given manifest file.

Return type
Manifest

model_config: ClassVar[ConfigDict] = {'arbitrary_types_allowed': True}

Configuration for the model, should be a dictionary conforming to [*ConfigDict*][pydantic.config.ConfigDict].

classmethod output_manifest_from_input_manifest(*input_manifest: Path | S3Path | Manifest*) → *Manifest*

Create Output manifest from input manifest file path, adds input files to output manifest configuration

Parameters
input_manifest (*Union[Path, S3Path, 'Manifest']*) – An S3 or regular path to an input_manifest object, or the input manifest object itself

Returns
output_manifest – The newly created output manifest

Return type
Manifest

serialize_filename(*filename: str | Path | S3Path | ManifestFilename | None*, *_info*) → str

Custom serializer for the manifest filename.

classmethod transform_filename(*raw_filename: str | ManifestFilename | None*) → *ManifestFilename* | None

Convert raw filename to ManifestFilename class if necessary.

classmethod transform_files(*raw_list: list[dict | str | Path | S3Path | ManifestFileRecord] | None*) → list[*ManifestFileRecord*]

Allow for the incoming files list to have varying types. Convert to a standardized list of ManifestFileStructure.

validate_checksums() → None

Validate checksums of listed files

write(*out_path: str | Path | S3Path*, *filename: str = None*) → Path | S3Path

Write a manifest file from a Manifest object (self).

Parameters

- **out_path** (*Union[str, Path, S3Path]*) – Directory path to write to (directory being used loosely to refer also to an S3 bucket path).
- **filename** (*str, Optional*) – must be a valid manifest filename. If not provided, the method uses the objects internal filename attribute. If that is not set, then a filename is automatically generated.

Returns
The path where the manifest file is written.

Return type
Union[Path, S3Path]

exception libera_utils.io.manifest.**ManifestError**

Generic exception related to manifest file handling

class libera_utils.io.manifest.**ManifestFileRecord**(*, *filename: str*, *checksum: str*)

Pydantic model for an individual data product file recorded within a manifest file.

Attributes

model_extra
Get extra fields set during validation.

model_fields_set
Returns the set of fields that have been explicitly set on this model instance.

Methods

copy(*[, include, exclude, update, deep])	Returns a copy of the model.
model_construct([_fields_set])	Creates a new instance of the *Model* class with validated data.
model_copy(*[, update, deep])	!!! abstract "Usage Documentation"
model_dump(*[, mode, include, exclude, ...])	!!! abstract "Usage Documentation"
model_dump_json(*[, indent, ensure_ascii, ...])	!!! abstract "Usage Documentation"
model_json_schema([by_alias, ref_template, ...])	Generates a JSON schema for a model class.
model_parametrized_name(params)	Compute the class name for parametrizations of generic classes.
model_post_init(context, /)	Override this method to perform additional initialization after *__init__* and *model_construct*.
model_rebuild(*[, force, raise_errors, ...])	Try to rebuild the pydantic-core schema for the model.
model_validate(obj, *[, strict, extra, ...])	Validate a pydantic model instance.
model_validate_json(json_data, *[, strict, ...])	!!! abstract "Usage Documentation"
model_validate_strings(obj, *[, strict, ...])	Validate the given object with string data against the Pydantic model.

construct
dict
from_orm
json
parse_file
parse_obj
parse_raw
schema
schema_json
update_forward_refs
validate

model_config: ClassVar[ConfigDict] = {}

Configuration for the model, should be a dictionary conforming to [*ConfigDict*][pydantic.config.ConfigDict].

libera_utils.io.manifest.**calculate_checksum**(*file: str | Path | S3Path*) → str

Compute the checksum of the given file.

libera_utils.io.manifest.**get_ulid_code**(*filename: str | Path | S3Path | ManifestFilename | None*) → ULID | None

Get ULID code from filename.

libera_utils.io.netcdf

Module containing utilities for writing Libera-conforming NetCDF4 data products

Functions

write_libera_data_product(...[, strict])	Write a Libera data product NetCDF4 file that conforms to data product definition requirements

libera_utils.io.netcdf.write_libera_data_product

libera_utils.io.netcdf.**write_libera_data_product**(*data_product_definition: str | CloudPath | Path, data: dict[str, ndarray[Any, dtype[_ScalarType_co]]], output_path: str | CloudPath | Path, time_variable: str, strict: bool = True*) → *LiberaDataProductFilename*

Write a Libera data product NetCDF4 file that conforms to data product definition requirements

Parameters

- **data_product_definition** (*str | PathType*) – Path to the data product definition against which to verify conformance
- **data** (*dict[str, NDarray]*) – Data mapping variable names to numpy data arrays
- **output_path** (*str | PathType*) – Base path (directory or S3 prefix) at which to write the product file
- **time_variable** (*str*) – Name of variable that indicates time. This is used to generate the start and end time for the filename.
- **strict** (*bool*) – Default True. Raises an exception if the final Dataset doesn't conform to the data product definition.

Returns

Filename object containing the full path to the written NetCDF4 data product file.

Return type

LiberaDataProductFilename

Classes

NetcdfEngine(value)	String enum class for our allowed NetCDF engines for xarray

libera_utils.io.netcdf.NetcdfEngine

class libera_utils.io.netcdf.**NetcdfEngine**(*value*)

Bases: StrEnum

String enum class for our allowed NetCDF engines for xarray

Methods

capitalize(/)	Return a capitalized version of the string.
casefold(/)	Return a version of the string suitable for caseless comparisons.
center(width[, fillchar])	Return a centered string of length width.
count(sub[, start[, end]])	Return the number of non-overlapping occurrences of substring sub in string S[start:end].
encode(/[, encoding, errors])	Encode the string using the codec registered for encoding.
endswith(suffix[, start[, end]])	Return True if S ends with the specified suffix, False otherwise.
expandtabs(/[, tabsize])	Return a copy where all tab characters are expanded using spaces.
find(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
format(*args, **kwargs)	Return a formatted version of S, using substitutions from args and kwargs.
format_map(mapping)	Return a formatted version of S, using substitutions from mapping.
index(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
isalnum(/)	Return True if the string is an alpha-numeric string, False otherwise.
isalpha(/)	Return True if the string is an alphabetic string, False otherwise.
isascii(/)	Return True if all characters in the string are ASCII, False otherwise.
isdecimal(/)	Return True if the string is a decimal string, False otherwise.
isdigit(/)	Return True if the string is a digit string, False otherwise.
isidentifier(/)	Return True if the string is a valid Python identifier, False otherwise.
islower(/)	Return True if the string is a lowercase string, False otherwise.
isnumeric(/)	Return True if the string is a numeric string, False otherwise.
isprintable(/)	Return True if the string is printable, False otherwise.
isspace(/)	Return True if the string is a whitespace string, False otherwise.
istitle(/)	Return True if the string is a title-cased string, False otherwise.
isupper(/)	Return True if the string is an uppercase string, False otherwise.
join(iterable, /)	Concatenate any number of strings.
ljust(width[, fillchar])	Return a left-justified string of length width.
lower(/)	Return a copy of the string converted to lowercase.
lstrip([chars])	Return a copy of the string with leading whitespace removed.
maketrans(x[, y, z])	Return a translation table usable for str.translate().
partition(sep, /)	Partition the string into three parts using the given separator.

continues on next page

Table 87 – continued from previous page

removeprefix(prefix, /)	Return a str with the given prefix string removed if present.
removesuffix(suffix, /)	Return a str with the given suffix string removed if present.
replace(old, new[, count])	Return a copy with all occurrences of substring old replaced by new.
rfind(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rindex(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rjust(width[, fillchar])	Return a right-justified string of length width.
rpartition(sep, /)	Partition the string into three parts using the given separator.
rsplit(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
rstrip([chars])	Return a copy of the string with trailing whitespace removed.
split(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
splitlines(/[, keepends])	Return a list of the lines in the string, breaking at line boundaries.
startswith(prefix[, start[, end]])	Return True if S starts with the specified prefix, False otherwise.
strip([chars])	Return a copy of the string with leading and trailing whitespace removed.
swapcase(/)	Convert uppercase characters to lowercase and lowercase characters to uppercase.
title(/)	Return a version of the string where each word is titlecased.
translate(table, /)	Replace each character in the string using the given translation table.
upper(/)	Return a copy of the string converted to uppercase.
zfill(width, /)	Pad a numeric string with zeros on the left, to fill a field of the given width.

__init__(**args*, ***kwds*)

Methods

encode(/[, encoding, errors])	Encode the string using the codec registered for encoding.
replace(old, new[, count])	Return a copy with all occurrences of substring old replaced by new.
split(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
rsplit(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
join(iterable, /)	Concatenate any number of strings.
capitalize(/)	Return a capitalized version of the string.

continues on next page

Table 88 – continued from previous page

casefold(/)	Return a version of the string suitable for caseless comparisons.
title(/)	Return a version of the string where each word is titlecased.
center(width[, fillchar])	Return a centered string of length width.
count(sub[, start[, end]])	Return the number of non-overlapping occurrences of substring sub in string S[start:end].
expandtabs(/[, tabsize])	Return a copy where all tab characters are expanded using spaces.
find(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
partition(sep, /)	Partition the string into three parts using the given separator.
index(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
ljust(width[, fillchar])	Return a left-justified string of length width.
lower(/)	Return a copy of the string converted to lowercase.
lstrip([chars])	Return a copy of the string with leading whitespace removed.
rfind(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rindex(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rjust(width[, fillchar])	Return a right-justified string of length width.
rstrip([chars])	Return a copy of the string with trailing whitespace removed.
rpartition(sep, /)	Partition the string into three parts using the given separator.
splitlines(/[, keepends])	Return a list of the lines in the string, breaking at line boundaries.
strip([chars])	Return a copy of the string with leading and trailing whitespace removed.
swapcase(/)	Convert uppercase characters to lowercase and lowercase characters to uppercase.
translate(table, /)	Replace each character in the string using the given translation table.
upper(/)	Return a copy of the string converted to uppercase.
startswith(prefix[, start[, end]])	Return True if S starts with the specified prefix, False otherwise.
endswith(suffix[, start[, end]])	Return True if S ends with the specified suffix, False otherwise.
removeprefix(prefix, /)	Return a str with the given prefix string removed if present.
removesuffix(suffix, /)	Return a str with the given suffix string removed if present.
isascii(/)	Return True if all characters in the string are ASCII, False otherwise.
islower(/)	Return True if the string is a lowercase string, False otherwise.
isupper(/)	Return True if the string is an uppercase string, False otherwise.

continues on next page

Table 88 – continued from previous page

istitle(/)	Return True if the string is a title-cased string, False otherwise.
isspace(/)	Return True if the string is a whitespace string, False otherwise.
isdecimal(/)	Return True if the string is a decimal string, False otherwise.
isdigit(/)	Return True if the string is a digit string, False otherwise.
isnumeric(/)	Return True if the string is a numeric string, False otherwise.
isalpha(/)	Return True if the string is an alphabetic string, False otherwise.
isalnum(/)	Return True if the string is an alpha-numeric string, False otherwise.
isidentifier(/)	Return True if the string is a valid Python identifier, False otherwise.
isprintable(/)	Return True if the string is printable, False otherwise.
zfill(width, /)	Pad a numeric string with zeros on the left, to fill a field of the given width.
format(*args, **kwargs)	Return a formatted version of S, using substitutions from args and kwargs.
format_map(mapping)	Return a formatted version of S, using substitutions from mapping.
maketrans(x[, y, z])	Return a translation table usable for str.translate().
get_from_config()	Retrieve the current netcdf engine config from the package configuration

Attributes

netcdf4
h5netcdf

capitalize(/)

Return a capitalized version of the string.

More specifically, make the first character have upper case and the rest lower case.

casefold(/)

Return a version of the string suitable for caseless comparisons.

center(*width*, *fillchar=' '*, /)

Return a centered string of length width.

Padding is done using the specified fill character (default is a space).

count(*sub*[, *start*[, *end*]]) → int

Return the number of non-overlapping occurrences of substring sub in string S[start:end]. Optional arguments start and end are interpreted as in slice notation.

encode(/, *encoding='utf-8'*, *errors='strict'*)

Encode the string using the codec registered for encoding.

encoding
The encoding in which to encode the string.

errors
The error handling scheme to use for encoding errors. The default is 'strict' meaning that encoding errors raise a UnicodeEncodeError. Other possible values are 'ignore', 'replace' and 'xmlcharrefreplace' as well as any other name registered with codecs.register_error that can handle UnicodeEncodeErrors.

endswith(*suffix*[, *start*[, *end*]]) → bool

Return True if S ends with the specified suffix, False otherwise. With optional start, test S beginning at that position. With optional end, stop comparing S at that position. suffix can also be a tuple of strings to try.

expandtabs(*/*, *tabsize=8*)

Return a copy where all tab characters are expanded using spaces.

If tabsize is not given, a tab size of 8 characters is assumed.

find(*sub*[, *start*[, *end*]]) → int

Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end]. Optional arguments start and end are interpreted as in slice notation.

Return -1 on failure.

format(**args*, ***kwargs*) → str

Return a formatted version of S, using substitutions from args and kwargs. The substitutions are identified by braces ('{' and '}').

format_map(*mapping*) → str

Return a formatted version of S, using substitutions from mapping. The substitutions are identified by braces ('{' and '}').

classmethod get_from_config() → Literal['netcdf4', 'h5netcdf']

Retrieve the current netcdf engine config from the package configuration

index(*sub*[, *start*[, *end*]]) → int

Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end]. Optional arguments start and end are interpreted as in slice notation.

Raises ValueError when the substring is not found.

isalnum(*/*)

Return True if the string is an alpha-numeric string, False otherwise.

A string is alpha-numeric if all characters in the string are alpha-numeric and there is at least one character in the string.

isalpha(*/*)

Return True if the string is an alphabetic string, False otherwise.

A string is alphabetic if all characters in the string are alphabetic and there is at least one character in the string.

isascii(*/*)

Return True if all characters in the string are ASCII, False otherwise.

ASCII characters have code points in the range U+0000-U+007F. Empty string is ASCII too.

isdecimal(/)

Return True if the string is a decimal string, False otherwise.

A string is a decimal string if all characters in the string are decimal and there is at least one character in the string.

isdigit(/)

Return True if the string is a digit string, False otherwise.

A string is a digit string if all characters in the string are digits and there is at least one character in the string.

isidentifier(/)

Return True if the string is a valid Python identifier, False otherwise.

Call keyword.iskeyword(s) to test whether string s is a reserved identifier, such as “def” or “class”.

islower(/)

Return True if the string is a lowercase string, False otherwise.

A string is lowercase if all cased characters in the string are lowercase and there is at least one cased character in the string.

isnumeric(/)

Return True if the string is a numeric string, False otherwise.

A string is numeric if all characters in the string are numeric and there is at least one character in the string.

isprintable(/)

Return True if the string is printable, False otherwise.

A string is printable if all of its characters are considered printable in repr() or if it is empty.

isspace(/)

Return True if the string is a whitespace string, False otherwise.

A string is whitespace if all characters in the string are whitespace and there is at least one character in the string.

istitle(/)

Return True if the string is a title-cased string, False otherwise.

In a title-cased string, upper- and title-case characters may only follow uncased characters and lowercase characters only cased ones.

isupper(/)

Return True if the string is an uppercase string, False otherwise.

A string is uppercase if all cased characters in the string are uppercase and there is at least one cased character in the string.

join(*iterable*, /)

Concatenate any number of strings.

The string whose method is called is inserted in between each given string. The result is returned as a new string.

Example: ‘.’.join([‘ab’, ‘pq’, ‘rs’]) -> ‘ab.pq.rs’

ljust(*width*, *fillchar=' '*, */*)

Return a left-justified string of length width.

Padding is done using the specified fill character (default is a space).

lower(*/*)

Return a copy of the string converted to lowercase.

lstrip(*chars=None*, */*)

Return a copy of the string with leading whitespace removed.

If chars is given and not None, remove characters in chars instead.

static maketrans(*x*, *y=<unrepresentable>*, *z=<unrepresentable>*, */*)

Return a translation table usable for str.translate().

If there is only one argument, it must be a dictionary mapping Unicode ordinals (integers) or characters to Unicode ordinals, strings or None. Character keys will be then converted to ordinals. If there are two arguments, they must be strings of equal length, and in the resulting dictionary, each character in x will be mapped to the character at the same position in y. If there is a third argument, it must be a string, whose characters will be mapped to None in the result.

partition(*sep*, */*)

Partition the string into three parts using the given separator.

This will search for the separator in the string. If the separator is found, returns a 3-tuple containing the part before the separator, the separator itself, and the part after it.

If the separator is not found, returns a 3-tuple containing the original string and two empty strings.

removeprefix(*prefix*, */*)

Return a str with the given prefix string removed if present.

If the string starts with the prefix string, return string[len(prefix):]. Otherwise, return a copy of the original string.

removesuffix(*suffix*, */*)

Return a str with the given suffix string removed if present.

If the string ends with the suffix string and that suffix is not empty, return string[:-len(suffix)]. Otherwise, return a copy of the original string.

replace(*old*, *new*, *count=-1*, */*)

Return a copy with all occurrences of substring old replaced by new.

count
Maximum number of occurrences to replace. -1 (the default value) means replace all occurrences.

If the optional argument count is given, only the first count occurrences are replaced.

rfind(*sub*[, *start*[, *end*]]) → int

Return the highest index in S where substring sub is found, such that sub is contained within S[start:end]. Optional arguments start and end are interpreted as in slice notation.

Return -1 on failure.

rindex(*sub*[, *start*[, *end*]]) → int

Return the highest index in S where substring sub is found, such that sub is contained within S[start:end]. Optional arguments start and end are interpreted as in slice notation.

Raises ValueError when the substring is not found.

rjust(*width*, *fillchar=' '*, */*)

Return a right-justified string of length width.

Padding is done using the specified fill character (default is a space).

rpartition(*sep*, */*)

Partition the string into three parts using the given separator.

This will search for the separator in the string, starting at the end. If the separator is found, returns a 3-tuple containing the part before the separator, the separator itself, and the part after it.

If the separator is not found, returns a 3-tuple containing two empty strings and the original string.

rsplit(*/*, *sep=None*, *maxsplit=-1*)

Return a list of the substrings in the string, using sep as the separator string.

sep
The separator used to split the string.

When set to None (the default value), will split on any whitespace character (including n r t f and spaces) and will discard empty strings from the result.

maxsplit
Maximum number of splits. -1 (the default value) means no limit.

Splitting starts at the end of the string and works to the front.

rstrip(*chars=None*, */*)

Return a copy of the string with trailing whitespace removed.

If chars is given and not None, remove characters in chars instead.

split(*/*, *sep=None*, *maxsplit=-1*)

Return a list of the substrings in the string, using sep as the separator string.

sep
The separator used to split the string.

When set to None (the default value), will split on any whitespace character (including n r t f and spaces) and will discard empty strings from the result.

maxsplit
Maximum number of splits. -1 (the default value) means no limit.

Splitting starts at the front of the string and works to the end.

Note, str.split() is mainly useful for data that has been intentionally delimited. With natural text that includes punctuation, consider using the regular expression module.

splitlines(*/*, *keepends=False*)

Return a list of the lines in the string, breaking at line boundaries.

Line breaks are not included in the resulting list unless keepends is given and true.

startswith(*prefix*[, *start*[, *end*]]) → bool

Return True if S starts with the specified prefix, False otherwise. With optional start, test S beginning at that position. With optional end, stop comparing S at that position. prefix can also be a tuple of strings to try.

strip(*chars=None*, */*)

Return a copy of the string with leading and trailing whitespace removed.

If chars is given and not None, remove characters in chars instead.

swapcase(/)

Convert uppercase characters to lowercase and lowercase characters to uppercase.

title(/)

Return a version of the string where each word is titlecased.

More specifically, words start with uppercased characters and all remaining cased characters have lower case.

translate(*table*, /)

Replace each character in the string using the given translation table.

table
Translation table, which must be a mapping of Unicode ordinals to Unicode ordinals, strings, or None.

The table must implement lookup/indexing via __getitem__, for instance a dictionary or list. If this operation raises LookupError, the character is left untouched. Characters mapped to None are deleted.

upper(/)

Return a copy of the string converted to uppercase.

zfill(*width*, /)

Pad a numeric string with zeros on the left, to fill a field of the given width.

The string is never truncated.

class libera_utils.io.netcdf.**NetcdfEngine**(*value*)

String enum class for our allowed NetCDF engines for xarray

Methods

capitalize(/)	Return a capitalized version of the string.
casefold(/)	Return a version of the string suitable for caseless comparisons.
center(width[, fillchar])	Return a centered string of length width.
count(sub[, start[, end]])	Return the number of non-overlapping occurrences of substring sub in string S[start:end].
encode(/[, encoding, errors])	Encode the string using the codec registered for encoding.
endswith(suffix[, start[, end]])	Return True if S ends with the specified suffix, False otherwise.
expandtabs(/[, tabsize])	Return a copy where all tab characters are expanded using spaces.
find(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
format(*args, **kwargs)	Return a formatted version of S, using substitutions from args and kwargs.
format_map(mapping)	Return a formatted version of S, using substitutions from mapping.
index(sub[, start[, end]])	Return the lowest index in S where substring sub is found, such that sub is contained within S[start:end].
isalnum(/)	Return True if the string is an alpha-numeric string, False otherwise.

continues on next page

Table 90 – continued from previous page

isalpha(/)	Return True if the string is an alphabetic string, False otherwise.
isascii(/)	Return True if all characters in the string are ASCII, False otherwise.
isdecimal(/)	Return True if the string is a decimal string, False otherwise.
isdigit(/)	Return True if the string is a digit string, False otherwise.
isidentifier(/)	Return True if the string is a valid Python identifier, False otherwise.
islower(/)	Return True if the string is a lowercase string, False otherwise.
isnumeric(/)	Return True if the string is a numeric string, False otherwise.
isprintable(/)	Return True if the string is printable, False otherwise.
isspace(/)	Return True if the string is a whitespace string, False otherwise.
istitle(/)	Return True if the string is a title-cased string, False otherwise.
isupper(/)	Return True if the string is an uppercase string, False otherwise.
join(iterable, /)	Concatenate any number of strings.
ljust(width[, fillchar])	Return a left-justified string of length width.
lower(/)	Return a copy of the string converted to lowercase.
lstrip([chars])	Return a copy of the string with leading whitespace removed.
maketrans(x[, y, z])	Return a translation table usable for str.translate().
partition(sep, /)	Partition the string into three parts using the given separator.
removeprefix(prefix, /)	Return a str with the given prefix string removed if present.
removesuffix(suffix, /)	Return a str with the given suffix string removed if present.
replace(old, new[, count])	Return a copy with all occurrences of substring old replaced by new.
rfind(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rindex(sub[, start[, end]])	Return the highest index in S where substring sub is found, such that sub is contained within S[start:end].
rjust(width[, fillchar])	Return a right-justified string of length width.
rpartition(sep, /)	Partition the string into three parts using the given separator.
rsplit(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
rstrip([chars])	Return a copy of the string with trailing whitespace removed.
split(/[, sep, maxsplit])	Return a list of the substrings in the string, using sep as the separator string.
splitlines(/[, keepends])	Return a list of the lines in the string, breaking at line boundaries.
startswith(prefix[, start[, end]])	Return True if S starts with the specified prefix, False otherwise.

continues on next page

Table 90 – continued from previous page

strip([chars])	Return a copy of the string with leading and trailing whitespace removed.
swapcase(/)	Convert uppercase characters to lowercase and lowercase characters to uppercase.
title(/)	Return a version of the string where each word is titlecased.
translate(table, /)	Replace each character in the string using the given translation table.
upper(/)	Return a copy of the string converted to uppercase.
zfill(width, /)	Pad a numeric string with zeros on the left, to fill a field of the given width.

classmethod get_from_config() → Literal['netcdf4', 'h5netcdf']

Retrieve the current netcdf engine config from the package configuration

libera_utils.io.netcdf.**write_libera_data_product**(*data_product_definition: str | CloudPath | Path*, *data: dict[str, ndarray[Any, dtype[_ScalarType_co]]]*, *output_path: str | CloudPath | Path*, *time_variable: str*, *strict: bool = True*) → *LiberaDataProductFilename*

Write a Libera data product NetCDF4 file that conforms to data product definition requirements

Parameters

- **data_product_definition** (*str | PathType*) – Path to the data product definition against which to verify conformance
- **data** (*dict[str, NDarray]*) – Data mapping variable names to numpy data arrays
- **output_path** (*str | PathType*) – Base path (directory or S3 prefix) at which to write the product file
- **time_variable** (*str*) – Name of variable that indicates time. This is used to generate the start and end time for the filename.
- **strict** (*bool*) – Default True. Raises an exception if the final Dataset doesn't conform to the data product definition.

Returns

Filename object containing the full path to the written NetCDF4 data product file.

Return type

LiberaDataProductFilename

libera_utils.io.product_definition

Data Product configuration and writing for Libera NetCDF4 data product files

Classes

LiberaDataProductDefinition(*, coordinates, ...)	Pydantic model for a Libera data product definition.
LiberaVariableDefinition(*, dtype, ...)	Pydantic model for a Libera variable definition.

libera_utils.io.product_definition.LiberaDataProductDefinition

class libera_utils.io.product_definition.**LiberaDataProductDefinition**(*, *coordinates: dict[str, LiberaVariableDefinition]*, *variables: dict[str, LiberaVariableDefinition]*, *attributes: dict[str, Any]*)

Bases: BaseModel

Pydantic model for a Libera data product definition.

Used for validating existing data product Datasets with helper methods for creating valid Datasets and DataArrays.

data_variables

A dictionary of variable names and their corresponding LiberaVariable objects, which contain metadata and data.

Type
dict[str, LiberaVariable]

product_metadata

The metadata associated with the data product, including dynamic metadata and spatio-temporal metadata.

Type
ProductMetadata | None

Attributes

dynamic_attributes
Return product-level attributes that are dynamically defined (null values) in the data product definition

model_extra
Get extra fields set during validation.

model_fields_set
Returns the set of fields that have been explicitly set on this model instance.

static_attributes
Return product-level attributes that are statically defined (have values) in the data product definition

Methods

check_dataset_conformance(dataset[, strict])	Check the conformance of a Dataset object against a DataProductDefinition
copy(*[, include, exclude, update, deep])	Returns a copy of the model.
create_conforming_dataset(data[, ...])	Create a Dataset from numpy arrays that is valid against the data product definition
enforce_dataset_conformance(dataset)	Analyze and update a Dataset to conform to the expectations of the DataProductDefinition
from_yaml(product_definition_filepath)	Create a DataProductDefinition from a Libera data product definition YAML file.
generate_data_product_filename(utc_start, ...)	Generate a standardized Libera data product filename.

continues on next page

Table 92 – continued from previous page

model_construct([_fields_set])	Creates a new instance of the *Model* class with validated data.
model_copy(*[, update, deep])	!!! abstract "Usage Documentation"
model_dump(*[, mode, include, exclude, ...])	!!! abstract "Usage Documentation"
model_dump_json(*[, indent, ensure_ascii, ...])	!!! abstract "Usage Documentation"
model_json_schema([by_alias, ref_template, ...])	Generates a JSON schema for a model class.
model_parametrized_name(params)	Compute the class name for parametrizations of generic classes.
model_post_init(context, /)	Override this method to perform additional initialization after *__init__* and *model_construct*.
model_rebuild(*[, force, raise_errors, ...])	Try to rebuild the pydantic-core schema for the model.
model_validate(obj, *[, strict, extra, ...])	Validate a pydantic model instance.
model_validate_json(json_data, *[, strict, ...])	!!! abstract "Usage Documentation"
model_validate_strings(obj, *[, strict, ...])	Validate the given object with string data against the Pydantic model.

construct
dict
from_orm
json
parse_file
parse_obj
parse_raw
schema
schema_json
update_forward_refs
validate

__init__(**data: *Any*) → None

Create a new model by parsing and validating input data from keyword arguments.

Raises [*ValidationError*][pydantic_core.ValidationError] if the input data cannot be validated to form a valid model.

self is explicitly positional-only to allow *self* as a field name.

Methods

check_dataset_conformance(dataset[, strict])	Check the conformance of a Dataset object against a DataProductDefinition
create_conforming_dataset(data[, ...])	Create a Dataset from numpy arrays that is valid against the data product definition
enforce_dataset_conformance(dataset)	Analyze and update a Dataset to conform to the expectations of the DataProductDefinition
from_yaml(product_definition_filepath)	Create a DataProductDefinition from a Libera data product definition YAML file.
generate_data_product_filename(utc_start, ...)	Generate a standardized Libera data product filename.

Attributes

dynamic_attributes	Return product-level attributes that are dynamically defined (null values) in the data product definition
model_computed_fields	
model_config	Configuration for the model, should be a dictionary conforming to [*ConfigDict*][pydantic.config.ConfigDict].
model_extra	Get extra fields set during validation.
model_fields	
model_fields_set	Returns the set of fields that have been explicitly set on this model instance.
static_attributes	Return product-level attributes that are statically defined (have values) in the data product definition
coordinates	
variables	
attributes	

_check_dataset_attrs(*dataset_attrs: dict[str, Any]*) → list[str]

Validate the product level attributes of a Dataset against the product definition

Static attributes must match exactly. Some special attributes have their values checked for validity.

Parameters
dataset_attrs (*dict[str, Any]*) – Dataset attributes to validate

Returns
List of error messages describing problems found. Empty list if no problems.

Return type
list[str]

static _get_static_project_attributes(*file_path=None*)

Loads project-wide consistent product-level attribute metadata from a YAML file.

These global attributes are expected on every Libera data product so we store them in a global config.

Parameters
file_path (*Path*) – The path to the global attribute metadata YAML file.

Returns
Dictionary of key-value pairs for static product attributes.

Return type
dict

classmethod _set_attributes(*raw_attributes: dict[str, Any]*) → dict[str, Any]

Validates product level attributes and adds requirements for globally consistent attributes

Parameters
raw_attributes (*dict[str, Any]*) – The attributes specification in the product definition.

Returns
The validated attributes dictionary, including standard defaults that we always require.

Return type
dict[str, Any]

check_dataset_conformance(*dataset: Dataset*, *strict: bool = True*) → list[str]

Check the conformance of a Dataset object against a DataProductDefinition

This method is responsible only for finding errors, not fixing them.

Parameters

- **dataset** (*Dataset*) – Dataset object to validate against expectations in the product configuration
- **strict** (*bool*) – Default True. Raises an exception for nonconformance.

Returns
List of error messages describing problems found. Empty list if no problems.

Return type
list[str]

create_conforming_dataset(*data: dict[str, ndarray]*, *user_product_attributes: dict[str, Any] | None = None*, *user_variable_attributes: dict[str, dict[str, Any]] | None = None*, *strict: bool = True*) → tuple[Dataset, list[str]]

Create a Dataset from numpy arrays that is valid against the data product definition

Parameters

- **data** (*dict[str, np.ndarray]*) – Dictionary of variable/coordinate data keyed by variable/coordinate name.
- **user_product_attributes** (*dict[str, Any] | None*) – *Algorithm developers should not need to use this kwarg.* Product level attributes for the data product. This allows the user to specify product level attributes that are required but not statically specified in the product definition (e.g. the algorithm version used to generate the product)
- **user_variable_attributes** (*dict[str, dict[str, Any]] | None*) – *Algorithm developers should not need to use this kwarg.* Per-variable attributes for each variable's DataArray. Key is variable name, value is an attributes dict. This allows the user to specify variable level attributes that are required but not statically defined in the product definition.
- **strict** (*bool*) – Default True. Raises an exception for nonconformance.

Returns
Tuple of (Dataset, error_messages) where error_messages contains any validation problems. Empty list if the dataset is fully valid.

Return type
tuple[Dataset, list[str]]

Notes

- We make no distinction between coordinate and data variable input data and assume that we can determine which is which based on coordinate/variable names the product definition.
- This method is not responsible for primary validation or error reporting. We call out to check_dataset_conformance at the end for that.

property dynamic_attributes

Return product-level attributes that are dynamically defined (null values) in the data product definition

These attributes are _required_ but are expected to be defined externally to the data product definition

enforce_dataset_conformance(*dataset: Dataset*) → tuple[Dataset, list[str]]

Analyze and update a Dataset to conform to the expectations of the DataProductDefinition

This method is for modifying an existing xarray Dataset. If you are creating a Dataset from scratch with numpy arrays, consider using *create_conforming_dataset* instead.

Parameters
dataset (*Dataset*) – Possibly non-compliant dataset

Returns
Tuple of (updated Dataset, error_messages) where error_messages contains any problems that could not be fixed. Empty list if all problems were fixed.

Return type
tuple[Dataset, list[str]]

Notes

- This method is responsible for trying (and possibly failing) to coerce a Dataset into a valid form with attributes and encodings. We use check_dataset_conformance to check for validation errors.

classmethod from_yaml(*product_definition_filepath: str | CloudPath | Path*)

Create a DataProductDefinition from a Libera data product definition YAML file.

Parameters
product_definition_filepath (*str | PathType*) – Path to YAML file with product and variable definitions

Returns
Configured instance with loaded metadata and optional data

Return type
DataProductDefinition

generate_data_product_filename(*utc_start: datetime*, *utc_end: datetime*) → *LiberaDataProductFilename*

Generate a standardized Libera data product filename.

Parameters

- **utc_start** (*datetime*) – Start time of data in the file
- **utc_end** (*datetime*) – End time of data in the file

Returns
Properly formatted filename object

Return type
LiberaDataProductFilename

model_config: ClassVar[ConfigDict] = {'frozen': True}

Configuration for the model, should be a dictionary conforming to [*ConfigDict*][pydantic.config.ConfigDict].

property model_extra: dict[str, Any] | None

Get extra fields set during validation.

Returns

A dictionary of extra fields, or *None* if *config.extra* is not set to *"allow"*.

property model_fields_set: set[str]

Returns the set of fields that have been explicitly set on this model instance.

Returns

A set of strings representing the fields that have been set,
i.e. that were not filled from defaults.

property static_attributes

Return product-level attributes that are statically defined (have values) in the data product definition

libera_utils.io.product_definition.LiberaVariableDefinition

class libera_utils.io.product_definition.**LiberaVariableDefinition**(*, *dtype: str*, *attributes: dict[str, ~typing.Any] = {}*, *dimensions: list[str] = []*, *encoding: dict = <factory>*)

Bases: BaseModel

Pydantic model for a Libera variable definition.

This model is the same for both data variables and coordinate variables

dtype

The data type of the variable's data array, specified as a string

Type

str

attributes

The attribute metadata for the variable, containing specific key value pairs for CF metadata compliance

Type

VariableAttributes

dimensions

A list of dimensions that the variable's data array references. These should be instances of LiberaDimension.

Type

list[LiberaDimension]

encoding

A dictionary specifying how the variable's data should be encoded when written to a NetCDF file.

Type

dict

Attributes

dynamic_attributes
Return attributes for a variable that are dynamically defined (null values) in the data product definition

model_extra
Get extra fields set during validation.

model_fields_set
Returns the set of fields that have been explicitly set on this model instance.

static_attributes
Return attributes for a variable that are statically defined (have values) in the data product definition

Methods

check_data_array_conformance(data_array, ...)	Validate variable data array based on product definition.
copy(*[, include, exclude, update, deep])	Returns a copy of the model.
create_conforming_data_array(data, variable_name)	Create a DataArray for a single variable that is valid against the data product definition.
enforce_data_array_conformance(data_array, ...)	Analyze and fix a DataArray to conform to variable specifications in data product definition
model_construct([_fields_set])	Creates a new instance of the *Model* class with validated data.
model_copy(*[, update, deep])	!!! abstract "Usage Documentation"
model_dump(*[, mode, include, exclude, ...])	!!! abstract "Usage Documentation"
model_dump_json(*[, indent, ensure_ascii, ...])	!!! abstract "Usage Documentation"
model_json_schema([by_alias, ref_template, ...])	Generates a JSON schema for a model class.
model_parametrized_name(params)	Compute the class name for parametrizations of generic classes.
model_post_init(context, /)	Override this method to perform additional initialization after *__init__* and *model_construct*.
model_rebuild(*[, force, raise_errors, ...])	Try to rebuild the pydantic-core schema for the model.
model_validate(obj, *[, strict, extra, ...])	Validate a pydantic model instance.
model_validate_json(json_data, *[, strict, ...])	!!! abstract "Usage Documentation"
model_validate_strings(obj, *[, strict, ...])	Validate the given object with string data against the Pydantic model.

construct
dict
from_orm
json
parse_file
parse_obj
parse_raw
schema
schema_json
update_forward_refs
validate

__init__(**data: *Any*) → None

Create a new model by parsing and validating input data from keyword arguments.

Raises [*ValidationError*][pydantic_core.ValidationError] if the input data cannot be validated to form a valid model.

self is explicitly positional-only to allow *self* as a field name.

Methods

check_data_array_conformance(data_array, ...)	Validate variable data array based on product definition.
create_conforming_data_array(data, variable_name)	Create a DataArray for a single variable that is valid against the data product definition.
enforce_data_array_conformance(data_array, ...)	Analyze and fix a DataArray to conform to variable specifications in data product definition

Attributes

dynamic_attributes	Return attributes for a variable that are dynamically defined (null values) in the data product definition
model_computed_fields	
model_config	Configuration for the model, should be a dictionary conforming to [*ConfigDict*][pydantic.config.ConfigDict].
model_extra	Get extra fields set during validation.
model_fields	
model_fields_set	Returns the set of fields that have been explicitly set on this model instance.
static_attributes	Return attributes for a variable that are statically defined (have values) in the data product definition
dtype	
attributes	
dimensions	
encoding	

_check_data_array_attributes(*data_array_attrs: dict[str, Any]*, *variable_name: str*) → list[str]

Validate the variable level attributes of a DataArray against the product definition

Attributes must match exactly

Parameters

- **data_array_attrs** (*dict[str, Any]*) – DataArray attributes to validate
- **variable_name** (*str*) – Name of the variable being checked (for error messages)

Returns

List of error messages describing problems found. Empty list if no problems.

Return type

list[str]

classmethod _set_encoding(*encoding: dict | None*)

Merge configured encoding with required defaults, issuing warnings on conflicts.

check_data_array_conformance(*data_array: DataArray*, *variable_name: str*) → list[str]

Validate variable data array based on product definition.

This does not verify that all required coordinate data exists on the DataArray. Dimensions lacking coordinates are treated as index dimensions. If coordinate data is later added to a Dataset under a dimension of the same name, the dimension will reference that coordinate data.

Parameters

- **data_array** (*DataArray*) – The data array to validate with this variable's metadata configuration.
- **variable_name** (*str*) – Name of the variable being checked (for error messages)

Returns
List of error messages describing problems found. Empty list if no problems.

Return type
list[str]

create_conforming_data_array(*data: ndarray*, *variable_name: str*, *user_variable_attributes: dict[str, Any] | None = None*) → DataArray

Create a DataArray for a single variable that is valid against the data product definition.

Coordinate data is not required. Dimensions that reference coordinate dimensions are created as index dimensions. If coordinate data is added later (e.g. to a Dataset), these dimensions will reference the coordinates.

Parameters

- **data** (*np.ndarray*) – Data for the variable DataArray.
- **variable_name** (*str*) – Name of the variable. Used for log messages and warnings.
- **user_variable_attributes** (*dict[str, Any] | None*) – *Algorithm developers should not need to use this kwarg.* Variable level attributes defined by the user. This allows a user to specify dynamic attributes that may be required by the definition but not statically defined in yaml.

Returns
A valid DataArray for the specified variable

Return type
DataArray

property dynamic_attributes

Return attributes for a variable that are dynamically defined (null values) in the data product definition

These attributes are _required_ but are expected to be defined externally to the data product definition

enforce_data_array_conformance(*data_array: DataArray*, *variable_name: str*) → tuple[DataArray, list[str]]

Analyze and fix a DataArray to conform to variable specifications in data product definition

Parameters

- **data_array** (*DataArray*) – The variable data array to analyze and update
- **variable_name** (*str*) – Name of the variable being enforced (for logging)

Returns
Tuple of (updated DataArray, error_messages) where error_messages contains any problems that could not be fixed. Empty list if all problems were fixed.

Return type
tuple[DataArray, list[str]]

`model_config: ClassVar[ConfigDict] = {'frozen': True}`

Configuration for the model, should be a dictionary conforming to [*ConfigDict*][pydantic.config.ConfigDict].

`property model_extra: dict[str, Any] | None`

Get extra fields set during validation.

Returns
A dictionary of extra fields, or *None* if *config.extra* is not set to *"allow"*.

`property model_fields_set: set[str]`

Returns the set of fields that have been explicitly set on this model instance.

Returns

A set of strings representing the fields that have been set,
i.e. that were not filled from defaults.

`property static_attributes`

Return attributes for a variable that are statically defined (have values) in the data product definition

`class libera_utils.io.product_definition.LiberaDataProductDefinition(*, coordinates: dict[str, LiberaVariableDefinition], variables: dict[str, LiberaVariableDefinition], attributes: dict[str, Any])`

Pydantic model for a Libera data product definition.

Used for validating existing data product Datasets with helper methods for creating valid Datasets and DataArrays.

`data_variables`

A dictionary of variable names and their corresponding LiberaVariable objects, which contain metadata and data.

Type
dict[str, LiberaVariable]

`product_metadata`

The metadata associated with the data product, including dynamic metadata and spatio-temporal metadata.

Type
ProductMetadata | None

Attributes

`dynamic_attributes`
Return product-level attributes that are dynamically defined (null values) in the data product definition

`model_extra`
Get extra fields set during validation.

model_fields_set
Returns the set of fields that have been explicitly set on this model instance.

static_attributes
Return product-level attributes that are statically defined (have values) in the data product definition

Methods

check_dataset_conformance(dataset[, strict])	Check the conformance of a Dataset object against a DataProductDefinition
copy(*[, include, exclude, update, deep])	Returns a copy of the model.
create_conforming_dataset(data[, ...])	Create a Dataset from numpy arrays that is valid against the data product definition
enforce_dataset_conformance(dataset)	Analyze and update a Dataset to conform to the expectations of the DataProductDefinition
from_yaml(product_definition_filepath)	Create a DataProductDefinition from a Libera data product definition YAML file.
generate_data_product_filename(utc_start, ...)	Generate a standardized Libera data product filename.
model_construct([_fields_set])	Creates a new instance of the *Model* class with validated data.
model_copy(*[, update, deep])	!!! abstract "Usage Documentation"
model_dump(*[, mode, include, exclude, ...])	!!! abstract "Usage Documentation"
model_dump_json(*[, indent, ensure_ascii, ...])	!!! abstract "Usage Documentation"
model_json_schema([by_alias, ref_template, ...])	Generates a JSON schema for a model class.
model_parametrized_name(params)	Compute the class name for parametrizations of generic classes.
model_post_init(context, /)	Override this method to perform additional initialization after *__init__* and *model_construct*.
model_rebuild(*[, force, raise_errors, ...])	Try to rebuild the pydantic-core schema for the model.
model_validate(obj, *[, strict, extra, ...])	Validate a pydantic model instance.
model_validate_json(json_data, *[, strict, ...])	!!! abstract "Usage Documentation"
model_validate_strings(obj, *[, strict, ...])	Validate the given object with string data against the Pydantic model.

construct
dict
from_orm
json
parse_file
parse_obj
parse_raw
schema
schema_json
update_forward_refs
validate

_check_dataset_attrs(*dataset_attrs: dict[str, Any]*) → list[str]

Validate the product level attributes of a Dataset against the product definition

Static attributes must match exactly. Some special attributes have their values checked for validity.

Parameters
dataset_attrs (*dict[str, Any]*) – Dataset attributes to validate

Returns
List of error messages describing problems found. Empty list if no problems.

Return type
list[str]

static _get_static_project_attributes(*file_path=None*)

Loads project-wide consistent product-level attribute metadata from a YAML file.

These global attributes are expected on every Libera data product so we store them in a global config.

Parameters
file_path (*Path*) – The path to the global attribute metadata YAML file.

Returns
Dictionary of key-value pairs for static product attributes.

Return type
dict

classmethod _set_attributes(*raw_attributes: dict[str, Any]*) → dict[str, Any]

Validates product level attributes and adds requirements for globally consistent attributes

Parameters
raw_attributes (*dict[str, Any]*) – The attributes specification in the product definition.

Returns
The validated attributes dictionary, including standard defaults that we always require.

Return type
dict[str, Any]

check_dataset_conformance(*dataset: Dataset*, *strict: bool = True*) → list[str]

Check the conformance of a Dataset object against a DataProductDefinition

This method is responsible only for finding errors, not fixing them.

Parameters

- **dataset** (*Dataset*) – Dataset object to validate against expectations in the product configuration
- **strict** (*bool*) – Default True. Raises an exception for nonconformance.

Returns
List of error messages describing problems found. Empty list if no problems.

Return type
list[str]

create_conforming_dataset(*data: dict[str, ndarray]*, *user_product_attributes: dict[str, Any] | None = None*, *user_variable_attributes: dict[str, dict[str, Any]] | None = None*, *strict: bool = True*) → tuple[Dataset, list[str]]

Create a Dataset from numpy arrays that is valid against the data product definition

Parameters

- **data** (*dict[str, np.ndarray]*) – Dictionary of variable/coordinate data keyed by variable/coordinate name.
- **user_product_attributes** (*dict[str, Any] | None*) – *Algorithm developers should not need to use this kwarg.* Product level attributes for the data product. This allows the user to specify product level attributes that are required but not statically specified in the product definition (e.g. the algorithm version used to generate the product)
- **user_variable_attributes** (*dict[str, dict[str, Any]] | None*) – *Algorithm developers should not need to use this kwarg.* Per-variable attributes for each variable's DataArray. Key is variable name, value is an attributes dict. This allows the user to specify variable level attributes that are required but not statically defined in the product definition.
- **strict** (*bool*) – Default True. Raises an exception for nonconformance.

Returns
Tuple of (Dataset, error_messages) where error_messages contains any validation problems. Empty list if the dataset is fully valid.

Return type
tuple[Dataset, list[str]]

Notes

- We make no distinction between coordinate and data variable input data and assume that we can determine which is which based on coordinate/variable names the product definition.
- This method is not responsible for primary validation or error reporting. We call out to check_dataset_conformance at the end for that.

property dynamic_attributes

Return product-level attributes that are dynamically defined (null values) in the data product definition

These attributes are _required_ but are expected to be defined externally to the data product definition

enforce_dataset_conformance(*dataset: Dataset*) → tuple[Dataset, list[str]]

Analyze and update a Dataset to conform to the expectations of the DataProductDefinition

This method is for modifying an existing xarray Dataset. If you are creating a Dataset from scratch with numpy arrays, consider using *create_conforming_dataset* instead.

Parameters
dataset (*Dataset*) – Possibly non-compliant dataset

Returns
Tuple of (updated Dataset, error_messages) where error_messages contains any problems that could not be fixed. Empty list if all problems were fixed.

Return type
tuple[Dataset, list[str]]

Notes

- This method is responsible for trying (and possibly failing) to coerce a Dataset into a valid form with attributes and encodings. We use check_dataset_conformance to check for validation errors.

classmethod from_yaml(*product_definition_filepath: str | CloudPath | Path*)

Create a DataProductDefinition from a Libera data product definition YAML file.

Parameters
product_definition_filepath (*str | PathType*) – Path to YAML file with product and variable definitions

Returns
Configured instance with loaded metadata and optional data

Return type
DataProductDefinition

generate_data_product_filename(*utc_start: datetime*, *utc_end: datetime*) → *LiberaDataProductFilename*

Generate a standardized Libera data product filename.

Parameters

- **utc_start** (*datetime*) – Start time of data in the file
- **utc_end** (*datetime*) – End time of data in the file

Returns
Properly formatted filename object

Return type
LiberaDataProductFilename

model_config: ClassVar[ConfigDict] = {'frozen': True}

Configuration for the model, should be a dictionary conforming to [*ConfigDict*][pydantic.config.ConfigDict].

property static_attributes

Return product-level attributes that are statically defined (have values) in the data product definition

class libera_utils.io.product_definition.**LiberaVariableDefinition**(*, *dtype: str*, *attributes: dict[str, ~typing.Any] = {}*, *dimensions: list[str] = []*, *encoding: dict = <factory>*)

Pydantic model for a Libera variable definition.

This model is the same for both data variables and coordinate variables

dtype

The data type of the variable's data array, specified as a string

Type
str

attributes

The attribute metadata for the variable, containing specific key value pairs for CF metadata compliance

Type
VariableAttributes

dimensions

A list of dimensions that the variable's data array references. These should be instances of LiberaDimension.

Type
list[LiberaDimension]

encoding

A dictionary specifying how the variable's data should be encoded when written to a NetCDF file.

Type
dict

Attributes

dynamic_attributes
Return attributes for a variable that are dynamically defined (null values) in the data product definition

model_extra
Get extra fields set during validation.

model_fields_set
Returns the set of fields that have been explicitly set on this model instance.

static_attributes
Return attributes for a variable that are statically defined (have values) in the data product definition

Methods

check_data_array_conformance(data_array, ...)	Validate variable data array based on product definition.
copy(*[, include, exclude, update, deep])	Returns a copy of the model.
create_conforming_data_array(data, variable_name)	Create a DataArray for a single variable that is valid against the data product definition.
enforce_data_array_conformance(data_array, ...)	Analyze and fix a DataArray to conform to variable specifications in data product definition
model_construct([_fields_set])	Creates a new instance of the *Model* class with validated data.
model_copy(*[, update, deep])	!!! abstract "Usage Documentation"
model_dump(*[, mode, include, exclude, ...])	!!! abstract "Usage Documentation"
model_dump_json(*[, indent, ensure_ascii, ...])	!!! abstract "Usage Documentation"
model_json_schema([by_alias, ref_template, ...])	Generates a JSON schema for a model class.
model_parametrized_name(params)	Compute the class name for parametrizations of generic classes.
model_post_init(context, /)	Override this method to perform additional initialization after *__init__* and *model_construct*.
model_rebuild(*[, force, raise_errors, ...])	Try to rebuild the pydantic-core schema for the model.
model_validate(obj, *[, strict, extra, ...])	Validate a pydantic model instance.
model_validate_json(json_data, *[, strict, ...])	!!! abstract "Usage Documentation"
model_validate_strings(obj, *[, strict, ...])	Validate the given object with string data against the Pydantic model.

construct
dict
from_orm
json
parse_file
parse_obj
parse_raw
schema
schema_json
update_forward_refs
validate

_check_data_array_attributes(*data_array_attrs: dict[str, Any]*, *variable_name: str*) → list[str]

Validate the variable level attributes of a DataArray against the product definition

Attributes must match exactly

Parameters

- **data_array_attrs** (*dict[str, Any]*) – DataArray attributes to validate
- **variable_name** (*str*) – Name of the variable being checked (for error messages)

Returns
List of error messages describing problems found. Empty list if no problems.

Return type
list[str]

classmethod _set_encoding(*encoding: dict | None*)

Merge configured encoding with required defaults, issuing warnings on conflicts.

check_data_array_conformance(*data_array: DataArray*, *variable_name: str*) → list[str]

Validate variable data array based on product definition.

This does not verify that all required coordinate data exists on the DataArray. Dimensions lacking coordinates are treated as index dimensions. If coordinate data is later added to a Dataset under a dimension of the same name, the dimension will reference that coordinate data.

Parameters

- **data_array** (*DataArray*) – The data array to validate with this variable's metadata configuration.
- **variable_name** (*str*) – Name of the variable being checked (for error messages)

Returns
List of error messages describing problems found. Empty list if no problems.

Return type
list[str]

create_conforming_data_array(*data: ndarray*, *variable_name: str*, *user_variable_attributes: dict[str, Any] | None = None*) → DataArray

Create a DataArray for a single variable that is valid against the data product definition.

Coordinate data is not required. Dimensions that reference coordinate dimensions are created as index dimensions. If coordinate data is added later (e.g. to a Dataset), these dimensions will reference the coordinates.

Parameters

- **data** (*np.ndarray*) – Data for the variable DataArray.
- **variable_name** (*str*) – Name of the variable. Used for log messages and warnings.
- **user_variable_attributes** (*dict[str, Any] | None*) – *Algorithm developers should not need to use this kwarg.* Variable level attributes defined by the user. This allows a user to specify dynamic attributes that may be required by the definition but not statically defined in yaml.

Returns
A valid DataArray for the specified variable

Return type
DataArray

property dynamic_attributes

Return attributes for a variable that are dynamically defined (null values) in the data product definition

These attributes are _required_ but are expected to be defined externally to the data product definition

enforce_data_array_conformance(*data_array: DataArray*, *variable_name: str*) → tuple[DataArray, list[str]]

Analyze and fix a DataArray to conform to variable specifications in data product definition

Parameters

- **data_array** (*DataArray*) – The variable data array to analyze and update
- **variable_name** (*str*) – Name of the variable being enforced (for logging)

Returns
Tuple of (updated DataArray, error_messages) where error_messages contains any problems that could not be fixed. Empty list if all problems were fixed.

Return type
tuple[DataArray, list[str]]

model_config: ClassVar[ConfigDict] = {'frozen': True}

Configuration for the model, should be a dictionary conforming to [*ConfigDict*][pydantic.config.ConfigDict].

property static_attributes

Return attributes for a variable that are statically defined (have values) in the data product definition

libera_utils.io.smart_open

Module for smart_open

Functions

is_gzip(path)	Determine if a string points to a gzip file.
is_s3(path)	Determine if a string points to an s3 location or not.
smart_copy_file(source_path, dest_path[, delete])	Copy function that can handle local files or files in an S3 bucket.
smart_open(path[, mode, enable_gzip])	Open function that can handle local files or files in an S3 bucket.

libera_utils.io.smart_open.is_gzip

libera_utils.io.smart_open.**is_gzip**(*path: str | Path | S3Path*)

Determine if a string points to a gzip file.

Parameters
path (*Union[str, Path, S3Path]*) – Path to check.

Return type
bool

libera_utils.io.smart_open.is_s3

libera_utils.io.smart_open.**is_s3**(*path: str | Path | S3Path*)

Determine if a string points to an s3 location or not.

Parameters
path (*Union[str, Path, S3Path]*) – Path to determine if it is and s3 location or not.

Return type
bool

libera_utils.io.smart_open.smart_copy_file

libera_utils.io.smart_open.**smart_copy_file**(*source_path: str | Path | S3Path*, *dest_path: str | Path | S3Path*, *delete: bool | None = False*)

Copy function that can handle local files or files in an S3 bucket. Returns the path to the newly created file as a Path or an S3Path, depending on the destination.

Parameters

- **source_path** (*Union[str, Path, S3Path]*) – Path to the source file to be copied. Files residing in an s3 bucket must begin with "s3://".
- **dest_path** (*Union[str, Path, S3Path]*) – Path to the Destination file to be copied to. Files residing in an s3 bucket must begin with "s3://".
- **delete** (*bool, optional*) – If true, deletes files copied from source (default = False)

Returns
The path to the newly created file

Return type
Path or S3Path

libera_utils.io.smart_open.smart_open

libera_utils.io.smart_open.**smart_open**(*path: str | Path | S3Path*, *mode: str | None = 'rb'*, *enable_gzip: bool | None = True*)

Open function that can handle local files or files in an S3 bucket. It also correctly handles gzip files determined by a **.gz* extension.

Parameters

- **path** (*Union[str, Path, S3Path]*) – Path to the file to be opened. Files residing in an s3 bucket must begin with "s3://".
- **mode** (*str, Optional*) – Optional string specifying the mode in which the file is opened. Defaults to 'rb'.

- **enable_gzip** (*bool, Optional*) – Flag to specify that **.gz* files should be opened as a *GzipFile* object. Setting this to False is useful when creating the md5sum of a **.gz* file. Defaults to True.

Return type
IO or gzip.GzipFile

libera_utils.io.smart_open.**_copy_local_to_local**(*source_path: str | Path*, *dest_path: str | Path*, *delete: bool | None = False*)

Copy a local source file to a local destination.

Parameters

- **source_path** (*Union[str, Path]*) – Path to the source file to be copied.
- **dest_path** (*Union[str, Path]*) – Path to the destination for the copied file.
- **delete** (*bool, Optional*) – If true, deletes files copied from source (default = False)

Returns
The path to the newly created file

Return type
Path

libera_utils.io.smart_open.**_copy_local_to_s3**(*source_path: str | Path*, *dest_path: str | S3Path*, *delete: bool | None = False*)

Copy a local file to an S3 object.

Parameters

- **source_path** (*Union[str, Path]*) – Path to the source file to be copied.
- **dest_path** (*Union[str, S3Path]*) – Path to the destination for the copied file. Files residing in an s3 bucket must begin with "s3://".
- **delete** (*bool, optional*) – If true, deletes files copied from source (default = False)

Returns
The path to the newly created file

Return type
S3Path

libera_utils.io.smart_open.**_copy_s3_to_local**(*source_path: str | S3Path*, *dest_path: str | Path*, *delete: bool | None = False*)

Copy an S3 object to a local file.

Parameters

- **source_path** (*Union[str, S3Path]*) – Path to the source file to be copied. Files residing in an s3 bucket must begin with "s3://".
- **dest_path** (*Union[str, Path]*) – Path to the destination for the copied file.
- **delete** (*bool, optional*) – If true, deletes files copied from source (default = False)

Returns
The path to the newly created file

Return type
Path

libera_utils.io.smart_open._copy_s3_to_s3(*source_path: str | S3Path*, *dest_path: str | S3Path*, *delete: bool | None = False*)

Copy an S3 object to a different S3 object.

Parameters

- **source_path** (*Union[str, S3Path]*) – Path to the source file to be copied. Files residing in an s3 bucket must begin with "s3://".
- **dest_path** (*Union[str, S3Path]*) – Path to the Destination file to be copied to. Files residing in an s3 bucket must begin with "s3://".
- **delete** (*bool, optional*) – If true, deletes files copied from source (default = False)

Returns
The path to the newly created file

Return type
S3Path

libera_utils.io.smart_open.is_gzip(*path: str | Path | S3Path*)

Determine if a string points to a gzip file.

Parameters
path (*Union[str, Path, S3Path]*) – Path to check.

Return type
bool

libera_utils.io.smart_open.is_s3(*path: str | Path | S3Path*)

Determine if a string points to an s3 location or not.

Parameters
path (*Union[str, Path, S3Path]*) – Path to determine if it is and s3 location or not.

Return type
bool

libera_utils.io.smart_open.smart_copy_file(*source_path: str | Path | S3Path*, *dest_path: str | Path | S3Path*, *delete: bool | None = False*)

Copy function that can handle local files or files in an S3 bucket. Returns the path to the newly created file as a Path or an S3Path, depending on the destination.

Parameters

- **source_path** (*Union[str, Path, S3Path]*) – Path to the source file to be copied. Files residing in an s3 bucket must begin with "s3://".
- **dest_path** (*Union[str, Path, S3Path]*) – Path to the Destination file to be copied to. Files residing in an s3 bucket must begin with "s3://".
- **delete** (*bool, optional*) – If true, deletes files copied from source (default = False)

Returns
The path to the newly created file

Return type
Path or S3Path

libera_utils.io.smart_open.smart_open(*path: str | Path | S3Path*, *mode: str | None = 'rb'*, *enable_gzip: bool | None = True*)

Open function that can handle local files or files in an S3 bucket. It also correctly handles gzip files determined by a **.gz* extension.

Parameters

- **path** (*Union[str, Path, S3Path]*) – Path to the file to be opened. Files residing in an s3 bucket must begin with "s3://".
- **mode** (*str, Optional*) – Optional string specifying the mode in which the file is opened. Defaults to 'rb'.
- **enable_gzip** (*bool, Optional*) – Flag to specify that **.gz* files should be opened as a *GzipFile* object. Setting this to False is useful when creating the md5sum of a **.gz* file. Defaults to True.

Return type
IO or gzip.GzipFile

3.1.7 libera_utils.kernel_maker

Module containing CLI tool for creating SPICE kernels from packets

Functions

azel_kernel_cli_handler(parsed_args)	Generate SPICE Az/El kernels from command line arguments.
from_args(input_data_files, ...[, ...])	Create a SPICE kernel from an input file and kernel data product type.
from_manifest(input_manifest, ...[, ...])	Generate SPICE kernels from a manifest file.
jpss_kernel_cli_handler(parsed_args)	Generate SPICE JPSS kernels from command line arguments.
make_kernel(config_file, output_kernel[, ...])	Create a SPICE kernel from a configuration file and input data.
preprocess_data(input_data_file, ...)	Preprocess kernel data to perform conversions and determine time range.

libera_utils.kernel_maker.azel_kernel_cli_handler

libera_utils.kernel_maker.**azel_kernel_cli_handler**(*parsed_args: Namespace*)

Generate SPICE Az/El kernels from command line arguments.

Parameters
parsed_args (*argparse.Namespace*) – Namespace of parsed CLI arguments.

Returns
Output manifest file containing one or more kernel files.

Return type
libera_utils.io.manifest.Manifest

libera_utils.kernel_maker.from_args

libera_utils.kernel_maker.**from_args**(*input_data_files: list[str | CloudPath | Path]*, *kernel_identifier: str | DataProductIdentifier*, *output_dir: str | CloudPath | Path*, *overwrite=False*, *append=False*, *verbose=False*) → CloudPath | Path

Create a SPICE kernel from an input file and kernel data product type.

Parameters

- **input_data_files** (*list[str, filenaming.PathType]*) – Input data files.

- **kernel_identifier** (*str | DataProductIdentifier*) – Data product identifier that is associated with a kernel.
- **output_dir** (*str | filenaming.PathType*) – Output location for the SPICE kernels and output manifest.
- **overwrite** (*bool*) – Option to overwrite any existing similar-named SPICE kernels.
- **append** (*bool*) – Option to append to any existing similar-named SPICE kernels. If multiple input files are provided with append=False, the first file will create a new kernel, and subsequent files will append to it.
- **verbose** (*bool*) – Option to log with extra verbosity.

Returns
Output kernel file path.

Return type
filenaming.PathType

libera_utils.kernel_maker.from_manifest

libera_utils.kernel_maker.**from_manifest**(*input_manifest: str | CloudPath | Path*, *data_product_identifiers: list[str]*, *output_dir: str | CloudPath | Path*, *overwrite=False*, *append=False*, *verbose=False*)

Generate SPICE kernels from a manifest file.

Parameters

- **input_manifest** (*str | filenaming.PathType*) – Input manifest file containing one or more input data files.
- **data_product_identifiers** (*list[str]*) – One or more SPICE kernel data product identifiers.
- **output_dir** (*str | filenaming.PathType*) – Output location for the SPICE kernels and output manifest.
- **overwrite** (*bool, optional*) – Option to overwrite any existing similar-named SPICE kernels.
- **append** (*bool, optional*) – Option to append to any existing similar-named SPICE kernels.
- **verbose** (*bool, optional*) – Option to log with extra verbosity.

Returns
Output manifest file containing one or more kernel files.

Return type
libera_utils.io.manifest.Manifest

libera_utils.kernel_maker.jpss_kernel_cli_handler

libera_utils.kernel_maker.**jpss_kernel_cli_handler**(*parsed_args: Namespace*)

Generate SPICE JPSS kernels from command line arguments.

Parameters
parsed_args (*argparse.Namespace*) – Namespace of parsed CLI arguments.

Returns
Output manifest file containing one or more kernel files.

Return type
libera_utils.io.manifest.Manifest

libera_utils.kernel_maker.make_kernel

libera_utils.kernel_maker.**make_kernel**(*config_file: str | Path*, *output_kernel: str | CloudPath | Path*, *input_data: str | Path | None = None*, *overwrite: bool = False*, *append: bool | int = False*) → CloudPath | Path

Create a SPICE kernel from a configuration file and input data.

Parameters

- **config_file** (*str | pathlib.Path*) – JSON configuration file defining how to create the kernel.
- **output_kernel** (*str | filenaming.PathType*) – Output directory or file to create the kernel. If a directory, the file name will be based on the config_file, but with the SPICE file extension.
- **input_data** (*str | filenaming.PathType or pd.DataFrame, optional*) – Input data file or object. Not required if defined within the config.
- **overwrite** (*bool*) – Option to overwrite an existing file.
- **append** (*bool | int*) – Option to append to an existing file. Anything truthy will be treated as True.

Returns
Output kernel file path

Return type
filenaming.PathType

libera_utils.kernel_maker.preprocess_data

libera_utils.kernel_maker.**preprocess_data**(*input_data_file: str | CloudPath | Path*, *nominal_time_field: str*, *pkt_time_fields: Sequence[str]*, *kernel_identifier:* DataProductIdentifier) → tuple[DataFrame, tuple[datetime, datetime]]

Preprocess kernel data to perform conversions and determine time range.

Parameters

- **input_data_file** (*str | filenaming.PathType*) – Input data file.
- **nominal_time_field** (*str*) – Name of the field to store the converted time field as.
- **pkt_time_fields** (*Sequence[str]*) – Names of the telemetry packet time fields used to convert the time.
- **kernel_identifier** (DataProductIdentifier) – The kernel type being generated (needed to determine which packet reader to use).

Returns

- *pd.DataFrame* – Loaded SPICE kernel data.
- *datetime.datetime, datetime.datetime* – The date time range of the data.

libera_utils.kernel_maker.**azel_kernel_cli_handler**(*parsed_args: Namespace*)

Generate SPICE Az/El kernels from command line arguments.

Parameters
parsed_args (*argparse.Namespace*) – Namespace of parsed CLI arguments.

Returns
Output manifest file containing one or more kernel files.

Return type
libera_utils.io.manifest.Manifest

libera_utils.kernel_maker.**from_args**(*input_data_files: list[str | CloudPath | Path]*, *kernel_identifier: str | DataProductIdentifier*, *output_dir: str | CloudPath | Path*, *overwrite=False*, *append=False*, *verbose=False*) → CloudPath | Path

Create a SPICE kernel from an input file and kernel data product type.

Parameters

- **input_data_files** (*list[str, filenaming.PathType]*) – Input data files.
- **kernel_identifier** (*str | DataProductIdentifier*) – Data product identifier that is associated with a kernel.
- **output_dir** (*str | filenaming.PathType*) – Output location for the SPICE kernels and output manifest.
- **overwrite** (*bool*) – Option to overwrite any existing similar-named SPICE kernels.
- **append** (*bool*) – Option to append to any existing similar-named SPICE kernels. If multiple input files are provided with append=False, the first file will create a new kernel, and subsequent files will append to it.
- **verbose** (*bool*) – Option to log with extra verbosity.

Returns
Output kernel file path.

Return type
filenaming.PathType

libera_utils.kernel_maker.**from_manifest**(*input_manifest: str | CloudPath | Path*, *data_product_identifiers: list[str]*, *output_dir: str | CloudPath | Path*, *overwrite=False*, *append=False*, *verbose=False*)

Generate SPICE kernels from a manifest file.

Parameters

- **input_manifest** (*str | filenaming.PathType*) – Input manifest file containing one or more input data files.
- **data_product_identifiers** (*list[str]*) – One or more SPICE kernel data product identifiers.
- **output_dir** (*str | filenaming.PathType*) – Output location for the SPICE kernels and output manifest.
- **overwrite** (*bool, optional*) – Option to overwrite any existing similar-named SPICE kernels.
- **append** (*bool, optional*) – Option to append to any existing similar-named SPICE kernels.
- **verbose** (*bool, optional*) – Option to log with extra verbosity.

Returns
Output manifest file containing one or more kernel files.

Return type
libera_utils.io.manifest.Manifest

libera_utils.kernel_maker.**jpss_kernel_cli_handler**(*parsed_args: Namespace*)

Generate SPICE JPSS kernels from command line arguments.

Parameters
parsed_args (*argparse.Namespace*) – Namespace of parsed CLI arguments.

Returns
Output manifest file containing one or more kernel files.

Return type
libera_utils.io.manifest.Manifest

libera_utils.kernel_maker.**make_kernel**(*config_file: str | Path*, *output_kernel: str | CloudPath | Path*, *input_data: str | Path | None = None*, *overwrite: bool = False*, *append: bool | int = False*) → CloudPath | Path

Create a SPICE kernel from a configuration file and input data.

Parameters

- **config_file** (*str | pathlib.Path*) – JSON configuration file defining how to create the kernel.
- **output_kernel** (*str | filenaming.PathType*) – Output directory or file to create the kernel. If a directory, the file name will be based on the config_file, but with the SPICE file extension.
- **input_data** (*str | filenaming.PathType or pd.DataFrame, optional*) – Input data file or object. Not required if defined within the config.
- **overwrite** (*bool*) – Option to overwrite an existing file.
- **append** (*bool | int*) – Option to append to an existing file. Anything truthy will be treated as True.

Returns
Output kernel file path

Return type
filenaming.PathType

libera_utils.kernel_maker.**preprocess_data**(*input_data_file: str | CloudPath | Path*, *nominal_time_field: str*, *pkt_time_fields: Sequence[str]*, *kernel_identifier: DataProductIdentifier*) → tuple[DataFrame, tuple[datetime, datetime]]

Preprocess kernel data to perform conversions and determine time range.

Parameters

- **input_data_file** (*str | filenaming.PathType*) – Input data file.
- **nominal_time_field** (*str*) – Name of the field to store the converted time field as.
- **pkt_time_fields** (*Sequence[str]*) – Names of the telemetry packet time fields used to convert the time.
- **kernel_identifier** (*DataProductIdentifier*) – The kernel type being generated (needed to determine which packet reader to use).

Returns

- *pd.DataFrame* – Loaded SPICE kernel data.

- *datetime.datetime, datetime.datetime* – The date time range of the data.

3.1.8 libera_utils.logutil

Logging utilities

Functions

configure_static_logging(config_file)	Configure logging based on a static logging configuration yaml file.
configure_task_logging(task_id, *[, ...])	Configure logging for a specific task (e.g. a processing algorithm).
flush_cloudwatch_logs()	Force flush of all cloudwatch logging handlers.

libera_utils.logutil.configure_static_logging

libera_utils.logutil.**configure_static_logging**(*config_file: str | Path | S3Path*)

Configure logging based on a static logging configuration yaml file.

The yaml is interpreted as a dict configuration. There is no ability to customize this logging configuration at runtime.

Parameters
config_file (*cloudpathlib.anypath.AnyPath or str*) – Location of config file.

See also

configure_task_logging
Runtime modifiable logging configuration.

libera_utils.logutil.configure_task_logging

libera_utils.logutil.**configure_task_logging**(*task_id: str*, *, *limit_debug_loggers: Iterable[str] | str | None = None*, *console_log_level: str | int = 20*, *console_log_json: bool = False*, *log_dir: str | Path | S3Path | None = None*, *cloudwatch_log_group: str | None = None*)

Configure logging for a specific task (e.g. a processing algorithm).

File-based logging is always done at the DEBUG level. Watchtower-based cloudwatch logging is always done at the DEBUG level. Console logging level defaults to INFO but can be set with console_log_level.

Examples

Example 1: The following will configure DEBUG console-only logging for anything in your script but all other loggers will be limited to INFO level.

`python configure_task_logging("my-script", limit_debug_loggers=("__main__",), console_log_level=logging.DEBUG) `

Example 2: This will allow all debug messages through from all loggers and sets up file-based logging and a custom cloudwatch log group. Also console messages will be logged in serialized JSON.

```python configure_task_logging(“my-script”,
```

console_log_level=logging.DEBUG, log_dir=Path("/tmp/my-script"), console_log_json=True, cloudwatch_log_group="custom-log-group")

```

**Parameters**

- **task_id** (*str*) – Unique identifier by which to name the log file and cloudwatch log stream.
- **limit_debug_loggers** (*Optional[Union[Iterable[str] | str]]*) – A list of logger name prefixes from which you want to allow debug messages (blocks debug from all others). For example, if you are working on a package called *my_app* and using module level logging, all your loggers will be named like *my_app.module_name.submodule_name*. By setting this to *(my_app,)*, all loggers that are named *my_app.** will propagate debug messages while preventing spammy debug messages from installed libraries like boto3. If this is empty or None, all debug messages will propagate. To use this in scripts, either leave it unset or use *limit_debug_loggers=("__main__,)*.
- **console_log_level** (*str or int, Optional*) – Log level for console logging. If not specified, defaults to INFO
- **console_log_json** (*bool, Optional*) – If True, console logs will be JSON formatted. This is suitable for setting up loggers in AWS services that are automatically monitored by cloudwatch on stdout and stderr (e.g. Lambda or Batch)
- **log_dir** (*str or Path or S3Path, Optional*) – Log directory, which may be a local or S3Path. Default is None and results in no file-based logging.
- **cloudwatch_log_group** (*str, Optional*) – Override optional environment variable log group name. Default is None and will result in falling back to the LIBERA_LOG_GROUP environment variable. If that is not set, no cloudwatch JSON logging will be configured.

### Notes

Even in the absence of cloudwatch JSON logging, all stdout/stderr messages generated by a Lambda will be logged to CloudWatch as string messages. Embedded JSON strings in log message text can still be queried in CloudWatch.

**See also**

*configure_static_logging*
Static logging configuration based on yaml file.

## libera_utils.logutil.flush_cloudwatch_logs

libera_utils.logutil.**flush_cloudwatch_logs**()

Force flush of all cloudwatch logging handlers.

If you are missing the last few log messages in a log stream, this may help get those logs ingested before the process shuts down the logging system.

**Return type**
None
```

Classes

JsonLogFormatter(*args[, ...])	Altered version of the CloudWatchLogFormatter provided in the watchtower library

libera_utils.logutil.JsonLogFormatter

class libera_utils.logutil.**JsonLogFormatter**(**args*, *add_log_record_attrs: tuple[str, ...] | None = None*, *add_asctime: bool = True*, ***kwargs*)

Bases: Formatter

Altered version of the CloudWatchLogFormatter provided in the watchtower library

Methods

converter([seconds])	Convert seconds since the Epoch to a time tuple expressing local time.
format(record)	Format log message to a string
formatException(ei)	Format and return the specified exception information as a string.
formatStack(stack_info)	This method is provided as an extension point for specialized formatting of stack information.
formatTime(record[, datefmt])	Return the creation time of the specified LogRecord as formatted text.
usesTime()	Check if the format uses the creation time of the record.

formatMessage

__init__(**args*, *add_log_record_attrs: tuple[str, ...] | None = None*, *add_asctime: bool = True*, ***kwargs*)

Parameters

- **add_log_record_attrs** (*Optional, tuple*) – Tuple of log record attributes to add to the resulting structured JSON structure that comes out of the logging formatter.
- **add_asctime** (*bool*) – If True, adds an ASCII (ISO 8601-like) timestamp to the log record. Default True.

Methods

format(record)	Format log message to a string

Attributes

default_msec_format
default_time_format

continues on next page

Table 106 – continued from previous page

format(*record: LogRecord*) → str

Format log message to a string

Parameters
record (*logging.LogRecord*) – Log record object containing the logged message, which may be a dict (Mapping) or a string

class libera_utils.logutil.**JsonLogFormatter**(**args*, *add_log_record_attrs: tuple[str, ...] | None = None*, *add_asctime: bool = True*, ***kwargs*)

Altered version of the CloudWatchLogFormatter provided in the watchtower library

Methods

converter([seconds])	Convert seconds since the Epoch to a time tuple expressing local time.
format(record)	Format log message to a string
formatException(ei)	Format and return the specified exception information as a string.
formatStack(stack_info)	This method is provided as an extension point for specialized formatting of stack information.
formatTime(record[, datefmt])	Return the creation time of the specified LogRecord as formatted text.
usesTime()	Check if the format uses the creation time of the record.

formatMessage

format(*record: LogRecord*) → str

Format log message to a string

Parameters
record (*logging.LogRecord*) – Log record object containing the logged message, which may be a dict (Mapping) or a string

libera_utils.logutil.**_json_serialize_default**(*o: Any*) → str

A standard 'default' json serializer function.

- Serializes datetime objects using their .isoformat() method.
- Serializes all other objects using repr().

libera_utils.logutil.**configure_static_logging**(*config_file: str | Path | S3Path*)

Configure logging based on a static logging configuration yaml file.

The yaml is interpreted as a dict configuration. There is no ability to customize this logging configuration at runtime.

Parameters
config_file (*cloudpathlib.anypath.AnyPath or str*) – Location of config file.

See also

configure_task_logging
Runtime modifiable logging configuration.

libera_utils.logutil.**configure_task_logging**(*task_id: str*, *, *limit_debug_loggers: Iterable[str] | str | None = None*, *console_log_level: str | int = 20*, *console_log_json: bool = False*, *log_dir: str | Path | S3Path | None = None*, *cloudwatch_log_group: str | None = None*)

Configure logging for a specific task (e.g. a processing algorithm).

File-based logging is always done at the DEBUG level. Watchtower-based cloudwatch logging is always done at the DEBUG level. Console logging level defaults to INFO but can be set with console_log_level.

Examples

Example 1: The following will configure DEBUG console-only logging for anything in your script but all other loggers will be limited to INFO level.

`python configure_task_logging("my-script", limit_debug_loggers=("__main__",), console_log_level=logging.DEBUG) `

Example 2: This will allow all debug messages through from all loggers and sets up file-based logging and a custom cloudwatch log group. Also console messages will be logged in serialized JSON.

```python configure_task_logging(“my-script”,

console_log_level=logging.DEBUG, log_dir=Path(“/tmp/my-script”), console_log_json=True, cloudwatch_log_group=”custom-log-group”)

```

Parameters

- **task_id** (*str*) – Unique identifier by which to name the log file and cloudwatch log stream.
- **limit_debug_loggers** (*Optional[Union[Iterable[str] | str]]*) – A list of logger name prefixes from which you want to allow debug messages (blocks debug from all others). For example, if you are working on a package called *my_app* and using module level logging, all your loggers will be named like *my_app.module_name.submodule_name*. By setting this to *(my_app,)*, all loggers that are named *my_app.** will propagate debug messages while preventing spammy debug messages from installed libraries like boto3. If this is empty or None, all debug messages will propagate. To use this in scripts, either leave it unset or use *limit_debug_loggers=(“__main__,)*.
- **console_log_level** (*str or int, Optional*) – Log level for console logging. If not specified, defaults to INFO
- **console_log_json** (*bool, Optional*) – If True, console logs will be JSON formatted. This is suitable for setting up loggers in AWS services that are automatically monitored by cloudwatch on stdout and stderr (e.g. Lambda or Batch)
- **log_dir** (*str or Path or S3Path, Optional*) – Log directory, which may be a local or S3Path. Default is None and results in no file-based logging.
- **cloudwatch_log_group** (*str, Optional*) – Override optional environment variable log group name. Default is None and will result in falling back to the LIBERA_LOG_GROUP environment variable. If that is not set, no cloudwatch JSON logging will be configured.

Notes

Even in the absence of cloudwatch JSON logging, all stdout/stderr messages generated by a Lambda will be logged to CloudWatch as string messages. Embedded JSON strings in log message text can still be queried in CloudWatch.

See also

configure_static_logging
Static logging configuration based on yaml file.

libera_utils.logutil.**flush_cloudwatch_logs**()

Force flush of all cloudwatch logging handlers.

If you are missing the last few log messages in a log stream, this may help get those logs ingested before the process shuts down the logging system.

Return type
None

3.1.9 libera_utils.packets

Module for reading packet data using Space Packet Parser

Functions

parse_packets_to_dataframe(...[, apid, ...])	Parse packets from files into a pandas DataFrame using Space Packet Parser v6.0.0rc3.
read_azel_packet_data(packet_data_filepaths)	Read Az/El packet data from a list of file paths.
read_sc_packet_data(packet_data_filepaths[, ...])	Read spacecraft packet data from a list of file paths.

libera_utils.packets.parse_packets_to_dataframe

libera_utils.packets.**parse_packets_to_dataframe**(*packet_definition: str | CloudPath | Path | XtcePacketDefinition*, *packet_data_filepaths: list[str | CloudPath | Path]*, *apid: int | None = None*, *skip_header_bytes: int = 0*) → DataFrame

Parse packets from files into a pandas DataFrame using Space Packet Parser v6.0.0rc3.

Parameters

- **packet_definition** (*str | PathType | XtcePacketDefinition*) – XTCE packet definition file path or pre-loaded XtcePacketDefinition object.
- **packet_data_filepaths** (*list[str | PathType]*) – List of filepaths to packet files.
- **apid** (*Optional[int]*) – Filter on APID so we don't get mismatches in case the parser finds multiple parsable packet definitions in the files. This can happen if the XTCE document contains definitions for multiple packet types and >1 of those packet types is present in the packet data files.
- **skip_header_bytes** (*int*) – Number of header bytes to skip when reading packet files. Default is 0.

Returns
pandas DataFrame containing parsed packet data.

Return type
pd.DataFrame

libera_utils.packets.read_azel_packet_data

libera_utils.packets.**read_azel_packet_data**(*packet_data_filepaths: list[str | CloudPath | Path]*, *apid: int = 1048*) → DataFrame

Read Az/El packet data from a list of file paths.

Parameters

packet_data_filepaths
[list[str | Path | CloudPath]]] The list of file paths to the raw packet data

apid
[int] Application Packet ID to filter for. Default is 1048 for Az/El sample packets.

:returns: **packet_data – The configured packet data as a pandas DataFrame with restructured samples.**
:rtype: pd.DataFrame

libera_utils.packets.read_sc_packet_data

libera_utils.packets.**read_sc_packet_data**(*packet_data_filepaths: list[str | CloudPath | Path]*, *apid: int = 11*) → DataFrame

Read spacecraft packet data from a list of file paths.

Parameters

- **packet_data_filepaths** (*list[str | PathType]*) – The list of file paths to the raw packet data
- **apid** (*int*) – Application Packet ID to filter for. Default is 11 for JPSS geolocation packets.

Returns
packet_data – The configured packet data as a pandas DataFrame.

Return type
pd.DataFrame

libera_utils.packets.**parse_packets_to_dataframe**(*packet_definition: str | CloudPath | Path | XtcePacketDefinition*, *packet_data_filepaths: list[str | CloudPath | Path]*, *apid: int | None = None*, *skip_header_bytes: int = 0*) → DataFrame

Parse packets from files into a pandas DataFrame using Space Packet Parser v6.0.0rc3.

Parameters

- **packet_definition** (*str | PathType | XtcePacketDefinition*) – XTCE packet definition file path or pre-loaded XtcePacketDefinition object.
- **packet_data_filepaths** (*list[str | PathType]*) – List of filepaths to packet files.
- **apid** (*Optional[int]*) – Filter on APID so we don't get mismatches in case the parser finds multiple parsable packet definitions in the files. This can happen if the XTCE document contains definitions for multiple packet types and >1 of those packet types is present in the packet data files.

- **skip_header_bytes** (*int*) – Number of header bytes to skip when reading packet files. Default is 0.

Returns
pandas DataFrame containing parsed packet data.

Return type
pd.DataFrame

libera_utils.packets.**read_azel_packet_data**(*packet_data_filepaths: list[str | CloudPath | Path]*, *apid: int = 1048*) → DataFrame

Read Az/El packet data from a list of file paths.

Parameters

packet_data_filepaths
[list[str | Path | CloudPath]]] The list of file paths to the raw packet data

apid
[int] Application Packet ID to filter for. Default is 1048 for Az/El sample packets.

:returns: **packet_data – The configured packet data as a pandas DataFrame with restructured samples.**
:rtype: pd.DataFrame

libera_utils.packets.**read_sc_packet_data**(*packet_data_filepaths: list[str | CloudPath | Path]*, *apid: int = 11*) → DataFrame

Read spacecraft packet data from a list of file paths.

Parameters

- **packet_data_filepaths** (*list[str | PathType]*) – The list of file paths to the raw packet data
- **apid** (*int*) – Application Packet ID to filter for. Default is 11 for JPSS geolocation packets.

Returns
packet_data – The configured packet data as a pandas DataFrame.

Return type
pd.DataFrame

3.1.10 libera_utils.quality_flags

Quality flag definitions

Classes

FlagBit(*args[, message])	Subclass of int to capture both an integer value and an accompanying message
LiberaFlag(value)	Subclass of Flag that add a method for decomposing a flag into its individual components and a property to return a list of all messages associated with a quality flag
LiberaQualityFlag(value)	TODO[LIBSDC-610]: Once these quality flags are well defined, write tests against them

libera_utils.quality_flags.FlagBit

class libera_utils.quality_flags.**FlagBit**(**args*, *message=None*, ***kwargs*)

Bases: int

Subclass of int to capture both an integer value and an accompanying message

Attributes

denominator
the denominator of a rational number in lowest terms

imag
the imaginary part of a complex number

numerator
the numerator of a rational number in lowest terms

real
the real part of a complex number

Methods

as_integer_ratio(/)	Return integer ratio.
bit_count(/)	Number of ones in the binary representation of the absolute value of self.
bit_length(/)	Number of bits necessary to represent self in binary.
conjugate	Returns self, the complex conjugate of any int.
from_bytes(/, bytes[, byteorder, signed])	Return the integer represented by the given array of bytes.
to_bytes(/[, length, byteorder, signed])	Return an array of bytes representing an integer.

__init__(**args*, ***kwargs*)

Methods

as_integer_ratio(/)	Return integer ratio.
bit_count(/)	Number of ones in the binary representation of the absolute value of self.
bit_length(/)	Number of bits necessary to represent self in binary.
conjugate	Returns self, the complex conjugate of any int.
from_bytes(/, bytes[, byteorder, signed])	Return the integer represented by the given array of bytes.
to_bytes(/[, length, byteorder, signed])	Return an array of bytes representing an integer.

Attributes

denominator	the denominator of a rational number in lowest terms
imag	the imaginary part of a complex number
numerator	the numerator of a rational number in lowest terms
real	the real part of a complex number

as_integer_ratio(/)

Return integer ratio.

Return a pair of integers, whose ratio is exactly equal to the original int and with a positive denominator.

```
>>> (10).as_integer_ratio()
(10, 1)
>>> (-10).as_integer_ratio()
(-10, 1)
>>> (0).as_integer_ratio()
(0, 1)
```

bit_count(/)

Number of ones in the binary representation of the absolute value of self.

Also known as the population count.

```
>>> bin(13)
'0b1101'
>>> (13).bit_count()
3
```

bit_length(/)

Number of bits necessary to represent self in binary.

```
>>> bin(37)
'0b100101'
>>> (37).bit_length()
6
```

conjugate()

Returns self, the complex conjugate of any int.

denominator

the denominator of a rational number in lowest terms

classmethod from_bytes(/, *bytes*, *byteorder='big'*, *, *signed=False*)

Return the integer represented by the given array of bytes.

bytes
: Holds the array of bytes to convert. The argument must either support the buffer protocol or be an iterable object producing bytes. Bytes and bytearray are examples of built-in objects that support the buffer protocol.

byteorder
: The byte order used to represent the integer. If byteorder is 'big', the most significant byte is at the beginning of the byte array. If byteorder is 'little', the most significant byte is at the end of the byte array. To request the native byte order of the host system, use `sys.byteorder' as the byte order value. Default is to use 'big'.

signed
: Indicates whether two's complement is used to represent the integer.

imag

the imaginary part of a complex number

numerator
: the numerator of a rational number in lowest terms

real
: the real part of a complex number

to_bytes(*/*, *length=1*, *byteorder='big'*, ***, *signed=False*)
: Return an array of bytes representing an integer.

  length
  : Length of bytes object to use. An OverflowError is raised if the integer is not representable with the given number of bytes. Default is length 1.

  byteorder
  : The byte order used to represent the integer. If byteorder is 'big', the most significant byte is at the beginning of the byte array. If byteorder is 'little', the most significant byte is at the end of the byte array. To request the native byte order of the host system, use `sys.byteorder' as the byte order value. Default is to use 'big'.

  signed
  : Determines whether two's complement is used to represent the integer. If signed is False and a negative integer is given, an OverflowError is raised.

libera_utils.quality_flags.LiberaFlag

class libera_utils.quality_flags.**LiberaFlag**(*value*)

Bases: Flag

Subclass of Flag that add a method for decomposing a flag into its individual components and a property to return a list of all messages associated with a quality flag

__init__(**args*, ***kwds*)

Methods

decompose()	Return the set of all set flags that form a subset of the queried flag value.

Attributes

summary	Summarize quality flag value

decompose()
: Return the set of all set flags that form a subset of the queried flag value. Note that this is not the minimum set of quality flags but rather a full set of all flags such that when they are ORed together, they produce *self.value*

  Returns
  : A tuple containing (members, not_covered) *members* is a list of flag values that are subsets of *value not_covered* is zero if the OR of members recreates *value*. Non-zero otherwise if bits are set in *value* that do not exist as named values in cls.

  Return type
  : tuple

property summary

Summarize quality flag value

Returns
(value, message_list) where value is the integer value of the quality flag and message list is a list of strings describing the quality flag bits which are set.

Return type
tuple

libera_utils.quality_flags.LiberaQualityFlag

class libera_utils.quality_flags.**LiberaQualityFlag**(*value*)

Bases: *LiberaFlag*

TODO[LIBSDC-610]: Once these quality flags are well defined, write tests against them

__init__(**args*, ***kwds*)

Methods

decompose()	Return the set of all set flags that form a subset of the queried flag value.

Attributes

summary	Summarize quality flag value
MISSING_DATA	

decompose()

Return the set of all set flags that form a subset of the queried flag value. Note that this is not the minimum set of quality flags but rather a full set of all flags such that when they are ORed together, they produce *self.value*

Returns
A tuple containing (members, not_covered) *members* is a list of flag values that are subsets of *value not_covered* is zero if the OR of members recreates *value*. Non-zero otherwise if bits are set in *value* that do not exist as named values in cls.

Return type
tuple

property summary

Summarize quality flag value

Returns
(value, message_list) where value is the integer value of the quality flag and message list is a list of strings describing the quality flag bits which are set.

Return type
tuple

class libera_utils.quality_flags.**FlagBit**(**args*, *message=None*, ***kwargs*)

Subclass of int to capture both an integer value and an accompanying message

Attributes

denominator
the denominator of a rational number in lowest terms

imag
the imaginary part of a complex number

numerator
the numerator of a rational number in lowest terms

real
the real part of a complex number

Methods

as_integer_ratio(/)	Return integer ratio.
bit_count(/)	Number of ones in the binary representation of the absolute value of self.
bit_length(/)	Number of bits necessary to represent self in binary.
conjugate	Returns self, the complex conjugate of any int.
from_bytes(/, bytes[, byteorder, signed])	Return the integer represented by the given array of bytes.
to_bytes(/[, length, byteorder, signed])	Return an array of bytes representing an integer.

class libera_utils.quality_flags.**LiberaFlag**(*value*)

Subclass of Flag that add a method for decomposing a flag into its individual components and a property to return a list of all messages associated with a quality flag

decompose()

Return the set of all set flags that form a subset of the queried flag value. Note that this is not the minimum set of quality flags but rather a full set of all flags such that when they are ORed together, they produce *self.value*

Returns
A tuple containing (members, not_covered) *members* is a list of flag values that are subsets of *value not_covered* is zero if the OR of members recreates *value*. Non-zero otherwise if bits are set in *value* that do not exist as named values in cls.

Return type
tuple

property summary

Summarize quality flag value

Returns
(value, message_list) where value is the integer value of the quality flag and message list is a list of strings describing the quality flag bits which are set.

Return type
tuple

class libera_utils.quality_flags.**LiberaQualityFlag**(*value*)

TODO[LIBSDC-610]: Once these quality flags are well defined, write tests against them

3.1.11 libera_utils.spice_utils

Modules for SPICE kernel creation, management, and usage

Functions

ensure_spice([f_py, time_kernels_only])	Before trying to understand this piece of code, read this: https://stackoverflow.com/questions/5929107/decorators-with-parameters/60832711#60832711
find_most_recent_naif_kernel(naif_base_url, ...)	Retrieves the name of the most recent kernel at NAIF.
ls_kernel_coverage(kernel_type[, verbose])	List time coverage of all furnished kernels of a given type
ls_kernels([verbose, log])	List all furnished spice kernels.
ls_spice_constants([verbose])	List all constants in the Spice constant pool

libera_utils.spice_utils.ensure_spice

libera_utils.spice_utils.**ensure_spice**(*f_py: Callable = None*, *time_kernels_only: bool = False*)

Before trying to understand this piece of code, read this: https://stackoverflow.com/questions/5929107/decorators-with-parameters/60832711#60832711

Decorator/wrapper that tries to ensure that a metakernel is furnished in as complete a way as possible.

Control flow overview:

1. **Try simply calling the wrapped function naively.**
   - SUCCESS? Great! We're done.
   - SpiceyError? Go to step 2.
2. **Furnish metakernel at SPICE_METAKERNEL**
   - SUCCESS? Great, return the original function again (so it can be re-run).
   - KeyError? Seems like SPICE_METAKERNEL isn't set, no problem. Go to step 3.

Usage:

Three ways to use this object

1. A decorator with no arguments

```
@ensure_spice
def my_spicey_func(a, b):
    pass
```

2. A decorator with parameters. This is useful if we only need the latest SCLK and LSK kernels for the function involved.

```
@ensure_spice(time_kernels_only=True)
def my_spicey_time_func(a, b):
    pass
```

3. An explicit wrapper function, providing a dynamically set value for parameters, e.g. time_kernels_only

```
wrapped = ensure_spice(spicey_func, time_kernels_only=True)
result = wrapped(*args, **kwargs)
```

Parameters

- **f_py** (*Callable*) – The function requiring SPICE that we are going to wrap if being used explicitly, Otherwise None, in which case ensure_spice is being used, not as a function wrapper (see l2a_processing.py) but as a true decorator without an explicit function argument.
- **time_kernels_only** (*bool, Optional*) – Specify that we only need to furnish time kernels (if SPICE_METAKERNEL is set, we still just furnish that metakernel and assume the time kernels are included.

Returns
Decorated function, with spice error handling

Return type
Callable

libera_utils.spice_utils.find_most_recent_naif_kernel

libera_utils.spice_utils.**find_most_recent_naif_kernel**(*naif_base_url: str*, *kernel_file_regex: str*, *allowed_attempts: int = 3*) → str

Retrieves the name of the most recent kernel at NAIF.

Parameters

- **naif_base_url** (*str*) – URL to search for filenames matching kernel_file_regex
- **kernel_file_regex** (*str*) – Regular expression to match filenames on the naif website
- **allowed_attempts** (*int, Optional*) – Number of allowed download times for naif page default = 3

Returns
Returns the file name of the latest kernel on the naif page (e.g., "naif0012.tls")

Return type
str

libera_utils.spice_utils.ls_kernel_coverage

libera_utils.spice_utils.**ls_kernel_coverage**(*kernel_type: str*, *verbose: bool = False*) → dict

List time coverage of all furnished kernels of a given type

Parameters

- **kernel_type** (*str*) – Either 'CK' or 'SPK'
- **verbose** (*bool*) – If True, print to stdout also

Returns
Key is filename, value is a list of tuples giving the start and end times in ET.

Return type
dict

libera_utils.spice_utils.ls_kernels

libera_utils.spice_utils.**ls_kernels**(*verbose: bool = False*, *log: bool = False*) → list

List all furnished spice kernels.

Parameters

- **verbose** (*bool*) – If True, print to stdout also

- **log** (*bool*) – Whether or not to log the current kernel pool (this gets called a lot)

Returns
A list of KernelFileRecord named tuples.

Return type
list

libera_utils.spice_utils.ls_spice_constants

libera_utils.spice_utils.**ls_spice_constants**(*verbose: bool = False*) → dict

List all constants in the Spice constant pool

Parameters
verbose – If true, print to stdout also

Returns
Dictionary of kernel constants

Return type
dict

Classes

KernelFileCache(kernel_url[, max_cache_age, ...])	Class for downloading, caching, and furnishing SPICE kernel files locally.
KernelFileRecord(kernel_type, file_name)	Tuple for keeping track of kernel files with default kernel_level
SpiceBody(value)	Enum containing SPICE IDs for ephemeris bodies that we use.
SpiceFrame(value)	Enum containing SPICE IDs for reference frames, possibly defined in the Frame Kernel (FK)
SpiceId(strid, numid)	Class that represents a unique identifier in the NAIF SPICE library
SpiceInstrument(value)	Enum containing SPICE IDs for instrument geometries configured in the Instrument Kernel (IK)

libera_utils.spice_utils.KernelFileCache

class libera_utils.spice_utils.**KernelFileCache**(*kernel_url: str, max_cache_age: timedelta = datetime.timedelta(days=1), fallback_kernel: Path = None*)

Bases: object

Class for downloading, caching, and furnishing SPICE kernel files locally.

It attempts to find a cached kernel file in the user's cache directory (OS-specific location). If that file is not there or is old, it attempts to download it from the specified location. If it is unable to do that, it can optionally read a fallback file included in the libera_utils package but this is not recommended.

Attributes

cache_dir
Property that calls out to get the proper local cache directory

kernel_basename
Base filename of the kernel.

kernel_path
Return the local path location of the kernel if it exists.

Methods

clear()	Remove cached kernel file
download_kernel(kernel_url[, allowed_attempts])	Downloads a kernel from a URL or an S3 location to the system cache location.
furnsh()	Furnish the cached kernel
is_cached([include_stale])	Check the cache directory for kernel file that is within cache age limit.

__init__(*kernel_url: str*, *max_cache_age: timedelta = datetime.timedelta(days=1)*, *fallback_kernel: Path = None*)

Create a new file cache. Downloading is done on first access of kernel_path if the file is not already cached. Fallback occurs only after failing to download. :param kernel_url: Location of kernel file as a URL or an S3Path :type kernel_url: str or cloudpathlib.S3Path :param max_cache_age: Length of time to tolerate stale kernels in the cache without forcing a redownload. :type max_cache_age: datetime.timedelta :param fallback_kernel: Path pointing to a fallback kernel location. May be None, which disallows a fallback. :type fallback_kernel: pathlib.Path

Methods

clear()	Remove cached kernel file
download_kernel(kernel_url[, allowed_attempts])	Downloads a kernel from a URL or an S3 location to the system cache location.
furnsh()	Furnish the cached kernel
is_cached([include_stale])	Check the cache directory for kernel file that is within cache age limit.

Attributes

cache_dir	Property that calls out to get the proper local cache directory
kernel_basename	Base filename of the kernel.
kernel_path	Return the local path location of the kernel if it exists.

property cache_dir

Property that calls out to get the proper local cache directory

Returns
Path to the proper local cache for the system.

Return type
pathlib.Path

clear()

Remove cached kernel file

download_kernel(*kernel_url: str*, *allowed_attempts: int = 3*) → Path

Downloads a kernel from a URL or an S3 location to the system cache location.

Parameters

- **kernel_url** (*str*) – Filename of kernel on NAIF site, as discovered by find_most_recent_naif_kernel
- **allowed_attempts** (*int, Optional*) – Number of allowed download times for naif kernel default = 3

Returns
Location of downloaded file

Return type
pathlib.Path

furnsh()

Furnish the cached kernel

is_cached(*include_stale: bool = False*) → bool

Check the cache directory for kernel file that is within cache age limit. If present, return True.

Parameters
include_stale (*bool*) – Default False. If True, results include kernel that are past the max age.

Returns
Returns True if kernel is present locally and within the age limit.

Return type
bool

property kernel_basename

Base filename of the kernel.

Return type
str

property kernel_path: Path

Return the local path location of the kernel if it exists. If not, try downloading it. If that fails, return the fallback kernel, if allowed.

libera_utils.spice_utils.KernelFileRecord

class libera_utils.spice_utils.**KernelFileRecord**(*kernel_type: str, file_name: str*)

Bases: NamedTuple

Tuple for keeping track of kernel files with default kernel_level

Methods

count(value, /)	Return number of occurrences of value.
index(value[, start, stop])	Return first index of value.

__init__(**args, **kwargs*)

Methods

count(value, /)	Return number of occurrences of value.
index(value[, start, stop])	Return first index of value.

Attributes

file_name	Alias for field number 1
kernel_type	Alias for field number 0

count(*value*, /)

Return number of occurrences of value.

file_name: str

Alias for field number 1

index(*value*, *start=0*, *stop=sys.maxsize*, /)

Return first index of value.

Raises ValueError if the value is not present.

kernel_type: str

Alias for field number 0

libera_utils.spice_utils.SpiceBody

class libera_utils.spice_utils.**SpiceBody**(*value*)

Bases: Enum

Enum containing SPICE IDs for ephemeris bodies that we use.

__init__(**args*, ***kwds*)

Attributes

JPSS
SSB
SUN
EARTH
EARTH_MOON_BARYCENTER

libera_utils.spice_utils.SpiceFrame

class libera_utils.spice_utils.**SpiceFrame**(*value*)

Bases: Enum

Enum containing SPICE IDs for reference frames, possibly defined in the Frame Kernel (FK)

__init__(**args*, ***kwds*)

Attributes

J2000
ITRF93
EARTH_FIXED

libera_utils.spice_utils.SpiceId

class libera_utils.spice_utils.**SpiceId**(*strid: str*, *numid: int*)

Bases: NamedTuple

Class that represents a unique identifier in the NAIF SPICE library

Methods

count(value, /)	Return number of occurrences of value.
index(value[, start, stop])	Return first index of value.

__init__(**args*, ***kwargs*)

Methods

count(value, /)	Return number of occurrences of value.
index(value[, start, stop])	Return first index of value.

Attributes

numid	Alias for field number 1
strid	Alias for field number 0

count(*value*, /)

Return number of occurrences of value.

index(*value*, *start=0*, *stop=sys.maxsize*, /)

Return first index of value.

Raises ValueError if the value is not present.

numid: int

Alias for field number 1

strid: str

Alias for field number 0

libera_utils.spice_utils.SpiceInstrument

class libera_utils.spice_utils.**SpiceInstrument**(*value*)

Bases: Enum

Enum containing SPICE IDs for instrument geometries configured in the Instrument Kernel (IK)

__init__(**args*, ***kwds*)

class libera_utils.spice_utils.**KernelFileCache**(*kernel_url: str*, *max_cache_age: timedelta = datetime.timedelta(days=1)*, *fallback_kernel: Path = None*)

Class for downloading, caching, and furnishing SPICE kernel files locally.

It attempts to find a cached kernel file in the user's cache directory (OS-specific location). If that file is not there or is old, it attempts to download it from the specified location. If it is unable to do that, it can optionally read a fallback file included in the libera_utils package but this is not recommended.

Attributes

cache_dir
Property that calls out to get the proper local cache directory

kernel_basename
Base filename of the kernel.

kernel_path
Return the local path location of the kernel if it exists.

Methods

clear()	Remove cached kernel file
download_kernel(kernel_url[, allowed_attempts])	Downloads a kernel from a URL or an S3 location to the system cache location.
furnsh()	Furnish the cached kernel
is_cached([include_stale])	Check the cache directory for kernel file that is within cache age limit.

property cache_dir

Property that calls out to get the proper local cache directory

Returns
Path to the proper local cache for the system.

Return type
pathlib.Path

clear()

Remove cached kernel file

download_kernel(*kernel_url: str*, *allowed_attempts: int = 3*) → Path

Downloads a kernel from a URL or an S3 location to the system cache location.

Parameters

- **kernel_url** (*str*) – Filename of kernel on NAIF site, as discovered by find_most_recent_naif_kernel

- **allowed_attempts** (*int, Optional*) – Number of allowed download times for naif kernel default = 3

Returns
Location of downloaded file

Return type
pathlib.Path

furnsh()
Furnish the cached kernel

is_cached(*include_stale: bool = False*) → bool
Check the cache directory for kernel file that is within cache age limit. If present, return True.

Parameters
include_stale (*bool*) – Default False. If True, results include kernel that are past the max age.

Returns
Returns True if kernel is present locally and within the age limit.

Return type
bool

property kernel_basename
Base filename of the kernel.

Return type
str

property kernel_path: Path
Return the local path location of the kernel if it exists. If not, try downloading it. If that fails, return the fallback kernel, if allowed.

class libera_utils.spice_utils.**KernelFileRecord**(*kernel_type: str, file_name: str*)
Tuple for keeping track of kernel files with default kernel_level

Methods

count(value, /)	Return number of occurrences of value.
index(value[, start, stop])	Return first index of value.

file_name: str
Alias for field number 1

kernel_type: str
Alias for field number 0

class libera_utils.spice_utils.**SpiceBody**(*value*)
Enum containing SPICE IDs for ephemeris bodies that we use.

class libera_utils.spice_utils.**SpiceFrame**(*value*)
Enum containing SPICE IDs for reference frames, possibly defined in the Frame Kernel (FK)

class libera_utils.spice_utils.**SpiceId**(*strid: str, numid: int*)
Class that represents a unique identifier in the NAIF SPICE library

Methods

count(value, /)	Return number of occurrences of value.
index(value[, start, stop])	Return first index of value.

numid: *int*

Alias for field number 1

strid: *str*

Alias for field number 0

class libera_utils.spice_utils.**SpiceInstrument**(*value*)

Enum containing SPICE IDs for instrument geometries configured in the Instrument Kernel (IK)

libera_utils.spice_utils.**ensure_spice**(*f_py: Callable = None*, *time_kernels_only: bool = False*)

Before trying to understand this piece of code, read this: https://stackoverflow.com/questions/5929107/decorators-with-parameters/60832711#60832711

Decorator/wrapper that tries to ensure that a metakernel is furnished in as complete a way as possible.

Control flow overview:

1. **Try simply calling the wrapped function naively.**
   - SUCCESS? Great! We're done.
   - SpiceyError? Go to step 2.
2. **Furnish metakernel at SPICE_METAKERNEL**
   - SUCCESS? Great, return the original function again (so it can be re-run).
   - KeyError? Seems like SPICE_METAKERNEL isn't set, no problem. Go to step 3.

Usage:

Three ways to use this object

1. A decorator with no arguments

```
@ensure_spice
def my_spicey_func(a, b):
    pass
```

2. A decorator with parameters. This is useful if we only need the latest SCLK and LSK kernels for the function involved.

```
@ensure_spice(time_kernels_only=True)
def my_spicey_time_func(a, b):
    pass
```

3. An explicit wrapper function, providing a dynamically set value for parameters, e.g. time_kernels_only

```
wrapped = ensure_spice(spicey_func, time_kernels_only=True)
result = wrapped(*args, **kwargs)
```

Parameters

- **f_py** (*Callable*) – The function requiring SPICE that we are going to wrap if being used explicitly, Otherwise None, in which case ensure_spice is being used, not as a function wrapper (see l2a_processing.py) but as a true decorator without an explicit function argument.
- **time_kernels_only** (*bool, Optional*) – Specify that we only need to furnish time kernels (if SPICE_METAKERNEL is set, we still just furnish that metakernel and assume the time kernels are included.

Returns
Decorated function, with spice error handling

Return type
Callable

libera_utils.spice_utils.**find_most_recent_naif_kernel**(*naif_base_url: str*, *kernel_file_regex: str*, *allowed_attempts: int = 3*) → str

Retrieves the name of the most recent kernel at NAIF.

Parameters

- **naif_base_url** (*str*) – URL to search for filenames matching kernel_file_regex
- **kernel_file_regex** (*str*) – Regular expression to match filenames on the naif website
- **allowed_attempts** (*int, Optional*) – Number of allowed download times for naif page default = 3

Returns
Returns the file name of the latest kernel on the naif page (e.g., "naif0012.tls")

Return type
str

libera_utils.spice_utils.**ls_kernel_coverage**(*kernel_type: str*, *verbose: bool = False*) → dict

List time coverage of all furnished kernels of a given type

Parameters

- **kernel_type** (*str*) – Either 'CK' or 'SPK'
- **verbose** (*bool*) – If True, print to stdout also

Returns
Key is filename, value is a list of tuples giving the start and end times in ET.

Return type
dict

libera_utils.spice_utils.**ls_kernels**(*verbose: bool = False*, *log: bool = False*) → list

List all furnished spice kernels.

Parameters

- **verbose** (*bool*) – If True, print to stdout also
- **log** (*bool*) – Whether or not to log the current kernel pool (this gets called a lot)

Returns
A list of KernelFileRecord named tuples.

Return type
list

libera_utils.spice_utils.**ls_spice_constants**(*verbose: bool = False*) → dict

List all constants in the Spice constant pool

Parameters
verbose – If true, print to stdout also

Returns
Dictionary of kernel constants

Return type
dict

3.1.12 libera_utils.time

Module for dealing with time and time conventions

Some convention for this module

1. Only decorate direct spiceypy wrapper functions with the ensure_spice decorator. They should directly call a spiceypy function.

2. All spiceypy wrapper functions should read as <spiceypyfunc>_wrapper. We really only use these to allow array inputs for spiceypy functions that aren't already vectorized in C and to wrap them in ensure_spice.

3. All functions should have robust type-hinting.

Functions

convert_cds_integer_to_datetime(satellite_time)	Helper function to convert a satellite time given as an CCSDS Day Segmented Time Code (CDS) form as 8 byte integer to a timezone aware datetime object
et2utc_wrapper(et, fmt, prec)	Convert ephemeris times to UTC ISO strings.
et_2_datetime(et)	Convert ephemeris time to a python datetime object by first converting it to a UTC timestamp.
et_2_timestamp(et[, fmt])	Convert ephemeris time to a custom formatted timestamp (default is lowercase version of ISO).
multipart_to_dt64(data[, day_field, ...])	Convert multipart time fields to a datetime64 time.
sce2s_wrapper(et)	Convert ephemeris times to SCLK string https://naif.jpl.nasa.gov/pub/naif/toolkit_docs/C/cspice/sce2s_c.html Decorated wrapper for spiceypy.sce2s that will automatically furnish the latest metakernel and retry if the first call raises an exception.
scs2e_wrapper(sclk_str)	Convert SCLK strings to ephemeris time.
utc2et_wrapper(iso_str)	Convert UTC ISO strings to ephemeris times.

libera_utils.time.convert_cds_integer_to_datetime

libera_utils.time.**convert_cds_integer_to_datetime**(*satellite_time: int*)

Helper function to convert a satellite time given as an CCSDS Day Segmented Time Code (CDS) form as 8 byte integer to a timezone aware datetime object

Parameters
satellite_time (*int*) – A 64-bit unsigned integer that represents CDS time

Returns
cds_time

Return type
datetime.datetime

libera_utils.time.et2utc_wrapper

libera_utils.time.**et2utc_wrapper**(*et: float | Collection[float] | ndarray*, *fmt: str*, *prec: int*) → str | Collection[str]

Convert ephemeris times to UTC ISO strings. https://naif.jpl.nasa.gov/pub/naif/toolkit_docs/C/cspice/et2utc_c.html Decorated wrapper for spiceypy.et2utc that will automatically furnish the latest metakernel and retry if the first call raises an exception.

Parameters

- **et** (*Union[float, Collection[float], numpy.ndarray]*) – The ephemeris time value to be converted to UTC.
- **fmt** (*str*) – Format string defines the format of the output time string. See CSPICE docs.
- **prec** (*int*) – Number of digits of precision for fractional seconds.

Returns
UTC time string(s)

Return type
Union[numpy.ndarray, str]

libera_utils.time.et_2_datetime

libera_utils.time.**et_2_datetime**(*et: float | Collection[float] | ndarray*) → datetime | ndarray

Convert ephemeris time to a python datetime object by first converting it to a UTC timestamp.

Parameters
et (*float or Collection or numpy.ndarray*) – Ephemeris times to be converted.

Returns
Object representation of ephemeris times.

Return type
datetime.datetime or numpy.ndarray

libera_utils.time.et_2_timestamp

libera_utils.time.**et_2_timestamp**(*et: float | Collection[float] | ndarray*, *fmt: str = '%Y%m%dT%H%M%S.%f'*) → str | Collection[str]

Convert ephemeris time to a custom formatted timestamp (default is lowercase version of ISO).

Parameters

- **et** (*Union[float, Collection[float], numpy.ndarray]*) – Ephemeris Time to be converted.
- **fmt** (*str, Optional*) – Format string as defined by the datetime.strftime() function.

Returns
Formatted timestamps

Return type
Union[str, Collection[str]]

libera_utils.time.multipart_to_dt64

libera_utils.time.**multipart_to_dt64**(*data: DataFrame | Dataset*, *day_field: str | None = None*, *ms_field: str | None = None*, *us_field: str | None = None*, *s_field: str | None = None*, *epoch: str = '1958-01-01'*)

Convert multipart time fields to a datetime64 time.

Parameters

- **data** (*pd.DataFrame | xr.Dataset*) – Any data structure containing the named subscript-able fields.
- **day_field** (*str | None, optional*) – Name of the day count field.
- **ms_field** (*str | None, optional*) – Name of the millisecond count field.
- **us_field** (*str | None, optional*) – Name of the microsecond count field.
- **s_field** (*str | None, optional*) – Name of the second count field.
- **epoch** (*str, optional*) – Date time string of the zero-offset epoch. Default="1958-01-01"

Returns
Pandas series of the datetime64 values.

Return type
pd.Series

libera_utils.time.sce2s_wrapper

libera_utils.time.**sce2s_wrapper**(*et: float | Collection[float] | ndarray*) → str | ndarray

Convert ephemeris times to SCLK string https://naif.jpl.nasa.gov/pub/naif/toolkit_docs/C/cspice/sce2s_c.html Decorated wrapper for spiceypy.sce2s that will automatically furnish the latest metakernel and retry if the first call raises an exception.

Parameters
et (*Union[float, Collection[float], numpy.ndarray]*) – Ephemeris time

Returns
SCLK string

Return type
Union[str, Collection[str]]

libera_utils.time.scs2e_wrapper

libera_utils.time.**scs2e_wrapper**(*sclk_str: str | Collection[str]*) → float | ndarray

Convert SCLK strings to ephemeris time. https://naif.jpl.nasa.gov/pub/naif/toolkit_docs/C/cspice/scs2e_c.html Decorated wrapper for spiceypy.scs2e that will automatically furnish the latest metakernel and retry if the first call raises an exception.

Parameters
sclk_str (*Union[str, Collection[str]]*) – Spacecraft clock string

Returns
Ephemeris time

Return type
Union[float, numpy.ndarray]

libera_utils.time.utc2et_wrapper

libera_utils.time.**utc2et_wrapper**(*iso_str: str | Collection[str]*) → float | ndarray

Convert UTC ISO strings to ephemeris times. https://naif.jpl.nasa.gov/pub/naif/toolkit_docs/C/cspice/utc2et_c.html Decorated wrapper for spiceypy.utc2et that will automatically furnish the latest metakernel and retry if the first call raises an exception.

Parameters
iso_str (*Union[str, Collection[str]]*) – The UTC to convert to ephemeris time

Returns
Ephemeris time

Return type
float or numpy.ndarray

libera_utils.time.**convert_cds_integer_to_datetime**(*satellite_time: int*)

Helper function to convert a satellite time given as an CCSDS Day Segmented Time Code (CDS) form as 8 byte integer to a timezone aware datetime object

Parameters
satellite_time (*int*) – A 64-bit unsigned integer that represents CDS time

Returns
cds_time

Return type
datetime.datetime

libera_utils.time.**et2utc_wrapper**(*et: float | Collection[float] | ndarray*, *fmt: str*, *prec: int*) → str | Collection[str]

Convert ephemeris times to UTC ISO strings. https://naif.jpl.nasa.gov/pub/naif/toolkit_docs/C/cspice/et2utc_c.html Decorated wrapper for spiceypy.et2utc that will automatically furnish the latest metakernel and retry if the first call raises an exception.

Parameters

- **et** (*Union[float, Collection[float], numpy.ndarray]*) – The ephemeris time value to be converted to UTC.
- **fmt** (*str*) – Format string defines the format of the output time string. See CSPICE docs.
- **prec** (*int*) – Number of digits of precision for fractional seconds.

Returns
UTC time string(s)

Return type
Union[numpy.ndarray, str]

libera_utils.time.**et_2_datetime**(*et: float | Collection[float] | ndarray*) → datetime | ndarray

Convert ephemeris time to a python datetime object by first converting it to a UTC timestamp.

Parameters
et (*float or Collection or numpy.ndarray*) – Ephemeris times to be converted.

Returns
Object representation of ephemeris times.

Return type
datetime.datetime or numpy.ndarray

libera_utils.time.**et_2_timestamp**(*et: float | Collection[float] | ndarray*, *fmt: str = '%Y%m%dT%H%M%S.%f'*) → str | Collection[str]

Convert ephemeris time to a custom formatted timestamp (default is lowercase version of ISO).

Parameters

- **et** (*Union[float, Collection[float], numpy.ndarray]*) – Ephemeris Time to be converted.
- **fmt** (*str, Optional*) – Format string as defined by the datetime.strftime() function.

Returns
Formatted timestamps

Return type
Union[str, Collection[str]]

libera_utils.time.**multipart_to_dt64**(*data: DataFrame | Dataset*, *day_field: str | None = None*, *ms_field: str | None = None*, *us_field: str | None = None*, *s_field: str | None = None*, *epoch: str = '1958-01-01'*)

Convert multipart time fields to a datetime64 time.

Parameters

- **data** (*pd.DataFrame | xr.Dataset*) – Any data structure containing the named subscript-able fields.
- **day_field** (*str | None, optional*) – Name of the day count field.
- **ms_field** (*str | None, optional*) – Name of the millisecond count field.
- **us_field** (*str | None, optional*) – Name of the microsecond count field.
- **s_field** (*str | None, optional*) – Name of the second count field.
- **epoch** (*str, optional*) – Date time string of the zero-offset epoch. Default="1958-01-01"

Returns
Pandas series of the datetime64 values.

Return type
pd.Series

libera_utils.time.**sce2s_wrapper**(*et: float | Collection[float] | ndarray*) → str | ndarray

Convert ephemeris times to SCLK string https://naif.jpl.nasa.gov/pub/naif/toolkit_docs/C/cspice/sce2s_c.html Decorated wrapper for spiceypy.sce2s that will automatically furnish the latest metakernel and retry if the first call raises an exception.

Parameters
et (*Union[float, Collection[float], numpy.ndarray]*) – Ephemeris time

Returns
SCLK string

Return type
Union[str, Collection[str]]

libera_utils.time.**scs2e_wrapper**(*sclk_str: str | Collection[str]*) → float | ndarray

Convert SCLK strings to ephemeris time. https://naif.jpl.nasa.gov/pub/naif/toolkit_docs/C/cspice/scs2e_c.html Decorated wrapper for spiceypy.scs2e that will automatically furnish the latest metakernel and retry if the first call raises an exception.

Parameters
sclk_str (*Union[str, Collection[str]]*) – Spacecraft clock string

Returns
Ephemeris time

Return type
Union[float, numpy.ndarray]

libera_utils.time.**utc2et_wrapper**(*iso_str: str | Collection[str]*) → float | ndarray

Convert UTC ISO strings to ephemeris times. https://naif.jpl.nasa.gov/pub/naif/toolkit_docs/C/cspice/utc2et_c.html Decorated wrapper for spiceypy.utc2et that will automatically furnish the latest metakernel and retry if the first call raises an exception.

Parameters
iso_str (*Union[str, Collection[str]]*) – The UTC to convert to ephemeris time

Returns
Ephemeris time

Return type
float or numpy.ndarray

3.1.13 libera_utils.version

Module for anything related to package versioning

Functions

version()	Get package version from metadata

libera_utils.version.version

libera_utils.version.**version**()

Get package version from metadata

libera_utils.version.**version**()

Get package version from metadata

CHAPTER

FOUR

VERSION CHANGES

4.1 5.0.0

- FEAT: Add `write_libera_data_product` to `netcdf.py` for one-shot product writes from numpy arrays
- BREAKING: Remove L1A `ProcessingStepIdentifier` (node) enum constants
- BREAKING: Refactor `DataProductDefinition` (previously `DataProductConfig`) and product definition validation
- MAINT: Remove `hdf.py` and associated tests

4.2 4.0.0

- BREAKING: Removed L0 file chunking support: The chunk_number parameter and related methods have been removed from L0 data product identifiers
- BREAKING: Refactored identifier classes: Moved DataProductIdentifier and ProcessingStepIdentifier from libera_utils.aws.constants to libera_utils.constants
- BREAKING: Removed deprecated classes: ProductName, CkObject, and SpkObject enums have been removed and consolidated into DataProductIdentifier
- BREAKING: SPICE kernel filename consolidation: Merged separate SPICE kernel filename classes into LiberaProductFilename
- Enhanced DataProductIdentifier: Now includes embedded metadata with processing level information and simplified lookup methods
- Enhanced ProcessingStepIdentifier: Now includes embedded product relationships and improved level derivation
- New DataLevel enum: Added structured processing level enumeration with built-in archive bucket name mapping
- Filename validation: Updated L0 filename patterns to remove chunk number validation
- Improved type safety: Better type hints and validation throughout identifier classes
- Updated documentation: Refreshed user documentation to reflect new file naming conventions
- Added the Libera APID values we care about for science processing to the LiberaApid enum class

4.3 3.2.4

- Changed the SPICE kernel creation API
- Added the SPICE kernel creation Docker image
- Added the Geolocation Tier-1 integration test case and data files

4.4 3.2.3

- Improved ECR authentication issues with the `libera-utils ecr-upload` CLI command for working with SSO authentication
- Updated S3 list tool to print by default
- Improved tooling of the aws constants module for both DataProductIdentifier and ProcessingStepIdentifier
- Added step function and policy names to ProcessingStepIdentifier
- Updated constants for L2 data products
- Updated step function trigger tool with url output and naming improvements
- Exposed useful L2 tools at highest level of the `libera_utils` package

4.5 3.2.2

- Added dimension handling for data product configuration
- Added internal DataArray data storage matching to known dimensions
- Added jenkins file for publishing to PyPI
- Added contact listserv for project
- Added json and yaml linting to the pre-commit hooks
- Added prettier configuration for formatting markdown, json, and yaml files
- Added internal Dataset object to the Configurable NetCDF-4 metadata object
- Added writing of output file for the Configurable NetCDF-4 data product object

4.6 3.2.1

- Added automatic formatting to the repository using `pre-commit` hooks
- Added a list-serv for ci/cd output as part of the Jenkins files
- BUGFIX: Fixed filenaming changes from 3.2.0 that affected the CDK deployment and usage

4.7 3.2.0

- Added pydantic models for a Configurable NetCDF-4 metadata object
- Includes static metadata that is populated and known by these tools
- Includes dynamic metadata that must be provided by user of these tools
- Changed the naming scheme of the SPICE data files to include SPK or CK

4.8 3.1.1

- BUGFIX: Testing improvement for mocking docker image building

4.9 3.1.0

- Added Curryer (lasp-curryer) library dependency for SPICE kernel creation and geolocation
- Added static SPICE kernels and configuration files for geolocating NOAA-20 / CERES
- BREAKING: Removed geolocation submodule, replaced by interfaces within Curryer

4.10 3.0.0

- BREAKING: Removed support for python version 3.9 and 3.10
- BREAKING: Updated the aws constants ProcessingStepIdentifier dump method to be to_str_with_chunk_number
- Improved the standardization of CLI commands to call cli_handler functions that wrap the main functionality

4.11 2.5.2

- Added the s3-utils cli interface with subcommands put, list, and cp for ease of use s3 interactions
- The cli subcommand `libera-utils s3-utils put` will upload a file to the correct S3 archive bucket given an algorithm
- The cli subcommand `libera-utils s3-utils list` will list all files in a given S3 archive bucket for a given algorithm
- The cli subcommand `libera-utils s3-utils cp` will copy a file from one S3 archive bucket to another location
- Added the Libera Archive bucket naming pattern to the aws constants module
- Improved type hinting for the smart_open module
- Improved readability of cli testing
- Removing testing against python 3.9 and 3.10

4.12 2.5.1

- BUGFIX: missing pyyaml dependency that prevented usage of logutil module

4.13 2.5.0

- Reimplement Manifest class as a pydantic model and integrate with dependent code
- BREAKING: Remove the deprecated add_file_to_manifest method from Manifest class
- BREAKING: Rename outpath to out_path on the Manifest class write method

4.14 2.4.5

- Add `--ecr-image-tags` option to `libera-utils ecr-upload` CLI for tagging remote algorithm images in an ECR
- Add validation and serialization methods to the `DataProductIdentifier` and `ProcessingStepProductIdentifier` enums

4.15 2.4.4

- Allow overriding the standard docker config.json file with a minimal file for ECR uploading to prevent ECR upload permission failures from cached docker login credentials from CDK deployments

4.16 2.4.3

- Changes to ecr_upload to support programmatic building and pushing of Docker images
- Remove DynamoDB docs (moved to `libera_cdk`)

4.17 2.4.2

- BREAKING: Remove the `AnyFilename` polymorphic class. Please use `AbstractValidFilename.from_file_path()`

4.18 2.4.1

- Updating requirements of methods to use keyword arguments rather than positional arguments
- Adding ProcessingStepIdentifier and DataProductIdentifier to the filenaming classes
- Updating ecr names to work with the completion checker testing in libera_cdk

4.19 2.4.0

- Add properties to filenaming classes to retrieve `data_product_id` and `processing_step_id`
- Add ProcessingStepIdentifier and DataProductIdentifier standardization to be used by downstream repos

4.20 2.3.1

- Fix os.path.join bug in filenaming module that broke mocked S3 paths and also fix typehinting

4.21 2.3.0

- Create CLI tools for AWS ECR image upload and Step Function triggering
- Update manifest filenames to use ULID instead of timestamp for unique identifiers
- Change logutil configure_task_logging to optionally log JSON to console
- Allow configure_task_logging to optionally propagate DEBUG messages from specific loggers
- Update documentation for how the database is used in the Libera project in DynamoDB
- Create tools for DynamoDB in AWS for .pds files (CONS and PDS)
- Replace the use of PostgreSQL with DynamoDB for the Libera project

4.22 2.2.0

- Add AnyFilename polymorphic class
- Change filename of all products to a LiberaDataProductFilename that inherits from AnyFilename
- Update filenaming convention to be all capital letters
- Improve API for manifest module
- Add prefixing to Filename classes for predictable archive paths
- Add prefixing for manifest files for predictable and navigable paths in s3 buckets
- Update git to include lfs and move test data to lfs
- Improve database manager including caching improvements
- Improve smart_copy_file and bug fixes to smart_open testing
- Refactoring and improving pds ingest for database entries and integration testing in CDK
- Added handling of construction records and pds files appropriately when ingesting
  - This includes reading a construction record and removing the pds file entry for the construction record itself
- Improved testing of pds ingest and pds file orm models to more accurately reflect use cases
- Added output manifest creation from input manifest to match timestamps in filenames of input and output manifests
- Refactored pds ingest to use AnyPath objects for handling file locations
- Added error handling to pds ingest

4.23 2.1.1

- Update dependency specification to speed up dependency resolution wrt botocore/urllib3
- Improve database initialization to work with libera_cdk changes
- Fix bug in Dockerfile that incorrectly set the default entrypoint
- Add preliminary instrument kernel

4.24 2.1.0

- Improve API to Manifest and Manifest.add_files
- Add manifest filename enforcement to Manifest class
- Update filenaming conventions for product filenames and SPICE kernels
- Allow adding an s3 bucket/prefix as a basepath for filenames

4.25 2.0.1

- Remove the extras dependency spec because of the way SQLAlchemy imports models

4.26 2.0.0

- Add filenaming classes
- Add manifest file class
- Add construction record parser
- Update DB schema to store construction records
- Update kernel generation CLI to use manifest file pattern
- Shift database and spice related libraries to extras (not installed by default)
- Add smart_copy_file function that can copy files to and from S3 and filesystem locations transparently
- Remove HDF-EOS5 filehandling code
- Add quality flag classes
- Change license to BSD3

4.27 1.0.0

- Stub out project structure
- Add build and release processes to readme
- Switch to Poetry for project dependency configuration and build management
- Add geolocation module
- Add tools in spiceutil module for caching SPICE kernels from NAIF
- Add missing unit testing coverage
- Add spice.md documentation on how the package uses and manages SPICE kernels
- Add database tooling, dev database, and ORM setup
- Add smart_open for opening local or S3 objects
- Add logging utility functions for setting up application logging

CHAPTER

FIVE

LIBERA SCIENCE DATA PROCESSING UTILITIES

Libera Utils is a package containing modules that are commonly used throughout the Libera Science Data Center codebase and processing algorithms. This package is published on PyPI to support our L2 algorithm developers with standardized code for interacting with our AWS resources and a consistent API for common tasks required of all developers.

5.1 Documentation

Documentation site, including full API listing: https://lasp-libera-sdc-libera-utils.readthedocs-hosted.com

Additional documentation helpful for Level 2 Algorithm Developers is also available in the Libera SDC Developer Guide. Please contact the Libera SDC Team at LASP for access to the Developer Guide.

5.2 Installation

```
pip install libera-utils
```

Other suffixed versions such as release candidate versions (version strings suffixed with `rc` followed by the candidate number, e.g. `1.2.3rc2`) may also be available but are likely to contain bugs.

CHAPTER

SIX

DEVELOPING LIBERA UTILS

Libera Utils is versioned formally using semantic versioning and released as new features are made available and bugs are fixed. You can find the complete release history on PyPI. Release candidate (rc) versions are also released in order for the SDC Team to test new functionality without breaking downstream code using generous dependency specifications.

We recommend pinning major and minor release versions (e.g. 2.2) as minor releases may contain minor breaking changes. Patch releases will be restricted to bug fixes that do not cause breaking changes to existing APIs.

PYTHON MODULE INDEX

INDEX

Symbols

A

B

C

D

E

F

G

I

J

K

L

M

N

O

P

R

S

T

U

V

W

Z