

Прикладне програмування в науках про Землю

Практична робота №4

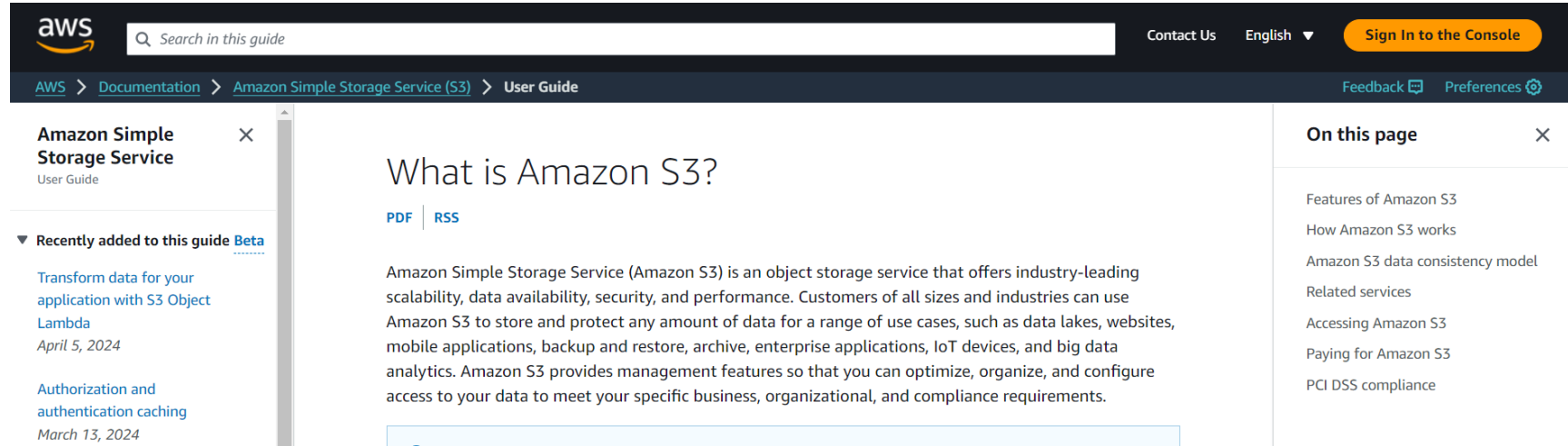
доц. Онищук В.І.
доц. Демидов В.К.
аспір. Охрімчук Р.Ю.

Хід роботи

1. додавання продуктів для роботи із Sentinel-2 та Landsat 5,7-9;
2. індексування супутникових знімків до ODC;
3. використання даних Sentinel-2 для геопросторового аналізу;
4. зробити звіт про виконану роботу в електронному вигляді.

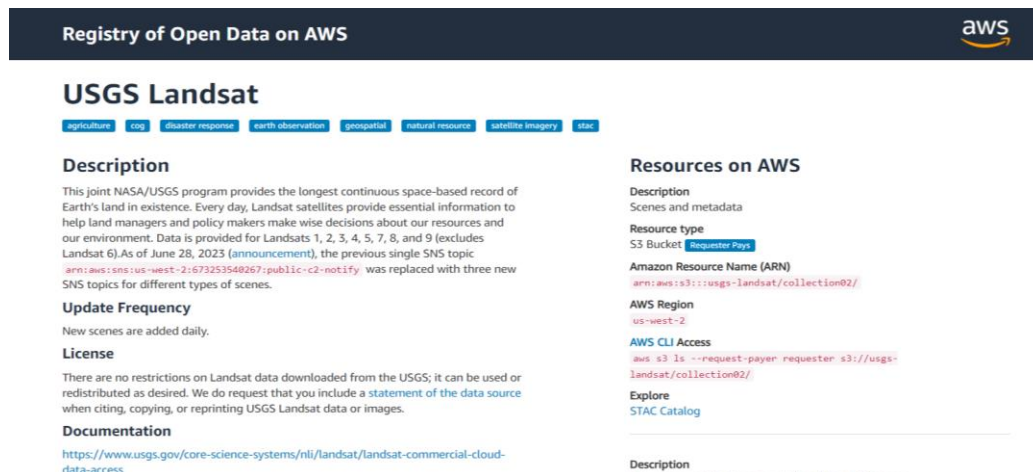
1. Додавання продуктів для роботи із Sentinel-2 та Landsat 5,7-9

1.1 Потрібно ознайомитися із S3 сервісом (Amazon Simple Storage Service) на офіційному електронному ресурсі AWS:

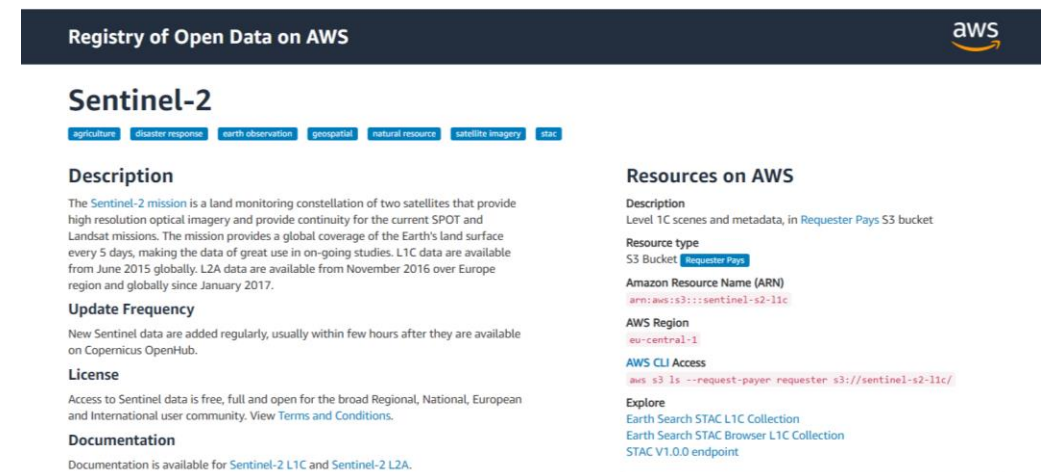


<https://docs.aws.amazon.com/AmazonS3/latest/userguide/Welcome.html>

1.2 Необхідно ознайомитися із сховищами, які містять супутникові знімки Landsat та Sentinel-2:



<https://registry.opendata.aws/usgs-landsat/> (<s3://usgs-landsat/collection02/>)



<https://registry.opendata.aws/sentinel-2/> (<s3://sentinel-s2-l1c/>)

1. Додавання продуктів для роботи із Sentinel-2 та Landsat 5,7-9

1.3 Необхідно додати до ODC продукти для Sentinel-2 та Landsat 5,7-9. Файли із метаданими продуктів знаходяться у папці – **products**.

Для додавання нових продуктів, потрібно:

- відкрити командний рядок **Anaconda Prompt**;
- активувати цільове середовище, яке попередньо налаштовано для роботи із **Open Data Cube**;
- наступним кроком є виконання команд у командному рядку **Anaconda Prompt**:
datacube product add s2_l2a_product.yaml (файл знаходиться у папці - **./configs**)
datacube product add landsat-c2l2-sr_products (файл знаходиться у папці - **./configs**)

Для перевірки, чи продукти були успішно додані до ODC, потрібно виконати команду:

datacube product list

Після виконання команди буде отримано список із усіма продуктами, які були додані до ODC. Якщо отриманий список включає у себе 5 продуктів, які вказані нижче, це буде підтвердженням успішного виконання даної операції

```
ls5_c2l2_sr    USGS Landsat 5 Collection 2 Level-2 Surface Reflectance
ls7_c2l2_sr    USGS Landsat 7 Collection 2 Level-2 Surface Reflectance
ls8_c2l2_sr    USGS Landsat 8 Collection 2 Level-2 Surface Reflectance
ls9_c2l2_sr    USGS Landsat 9 Collection 2 Level-2 Surface Reflectance
s2_l2a         Sentinel-2a and Sentinel-2b imagery, processed to Level 2A (Surface Reflectance) and converted to Cloud Optimized GeoTIFFs
```

2. Індексуння супутникових знімків до ODC

2.1 Потрібно встановити бібліотеку – **odc-apps-dc-tools**, яка містить набір інструментів, розроблених для роботи з Open Data Cube (ODC). Для інсталяції даної бібліотеки потрібно виконати команду у командному рядку Anaconda Prompt:

pip install odc-apps-dc-tools

2.2 Необхідно ознайомитися із роботою інструменту – **stac-to-dc**. Дана команда дозволяє користувачам автоматизувати процес завантаження і індексації даних з каталогів STAC (SpatioTemporal Asset Catalog) до локальної або хмарної інфраструктури ODC.

Основі параметри, які **stac-to-dc** приймає на вхід:

--catalog-href

Опис: URL каталогу STAC для імпорту даних.

Формат: **--catalog-href**=[URL]

Приклад: **--catalog-href**=https://earth-search.aws.element84.com/v1/

--collections

Опис: Назва або назви колекцій з каталогу STAC для імпорту.

Формат: **--collections**=[collection_name]

Приклад: **--collections**=sentinel-2-l2a

--bbox

Опис: Вказує географічну область для імпорту в форматі прямокутника.

Формат: **--bbox**=[xmin,ymin,xmax,ymax]

Приклад: **--bbox**=22.13,44.38,40.28,52.38

--datetime

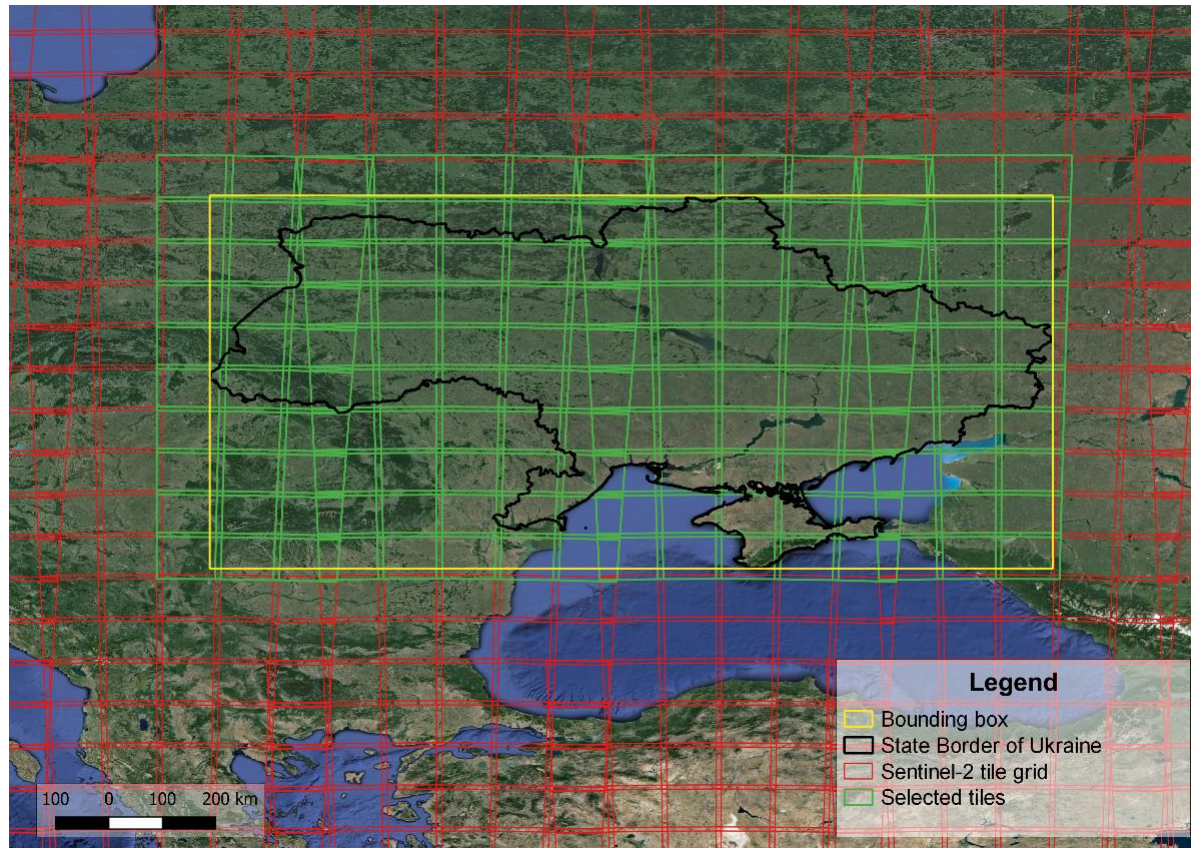
Опис: Часовий діапазон для імпорту даних.

Формат: **--datetime**=[YYYY-MM-DD/YYYY-MM-DD]

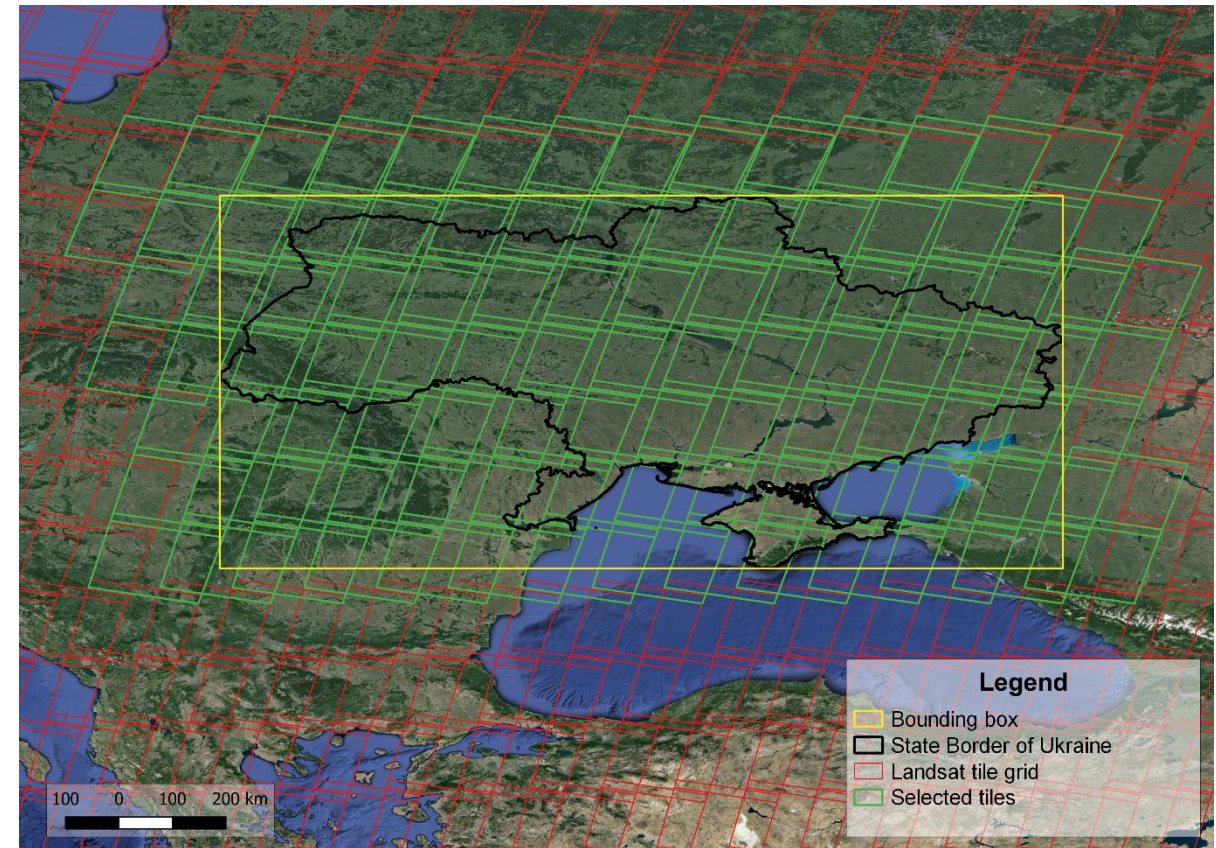
Приклад: **--datetime**=2024-01-01/2024-04-30

2. Індексунання супутникових знімків до ODC

Для кращого розуміння матеріалу, додаю візуалізацію того, як працює параметр **--bbox** у випадку, коли потрібно зробити індексацію сцен супутникових знімків, що покривають територію України. У даному випадку параметр **bbox** встановлено як 22.13,44.38,40.28,52.38, оскільки це координати полігону (показаного жовтим кольором), що охоплює територію України. Зеленим кольором показані тайли, які будуть додаватися до ODC. Сітку, яка схематично показує, яким чином тайли покривають земну поверхню (яка представлена у географічній системі координат - EPSG:4326), позначена червоним кольором.



Sentinel-2



Landsat 5, 7-9

2. Індексуння супутникових знімків до ODC

2.3 Потрібно виконати індексування даних **Sentinel-2** з S3 бакету за допомогою команди **stac-to-dc**, для цього потрібно виконати дану команду:

```
stac-to-dc --catalog-href=https://earth-search.aws.element84.com/v1/ --bbox=22.13,44.38,40.28,52.38 --collections=sentinel-2-l2a --datetime=2024-01-01/2024-05-31
```

(Дана команда додасть до продукту **s2_l2a** датасети, які містять метадані для всіх доступних у сховищі сцен супутникових знімків, що покривають територію України у період з **2024-01-01** по **2024-05-31**. Ці дані можна отримати без акаунту AWS)

Також додатково надаю приклад команди для індексування даних для продуктів **Landsat**:

```
stac-to-dc --catalog-href=https://landsatlook.usgs.gov/stac-server/ --rewrite-assets=https://landsatlook.usgs.gov/data/,s3://usgs-landsat/ --bbox=22.13,44.38,40.28,52.38 --collections=landsat-c2l2-sr --datetime=2024-01-01/2024-04-30
```

(Для отримання даних **Landsat** потрібно мати акаунт AWS, який буде дозволяти безкоштовно викачувати протягом місяця **100 Gb** даних із S3 сховищ)

Якщо датасети були додані успішно, буде отримано повідомлення, яке містить інформацію про кількість успішних операцій, як це наведено у прикладі нижче:

```
04/09/2024 07:16:57: WARNING: Dataset already exists, not indexing
04/09/2024 07:16:57: WARNING: Dataset already exists, not indexing
04/09/2024 07:16:57: WARNING: Dataset already exists, not indexing
04/09/2024 07:16:57: WARNING: Dataset already exists, not indexing
Added 350 Datasets, failed 0 Datasets, skipped 6759 Datasets
```

2. Індексуння супутникових знімків до ODC

При встановленні бібліотеки **odc-apps-dc-tools** можливе виникнення конфлікту версій бібліотеки **pathlib**, що проявлятиметься у вигляді помилки **ImportError** при індексуванні даних:

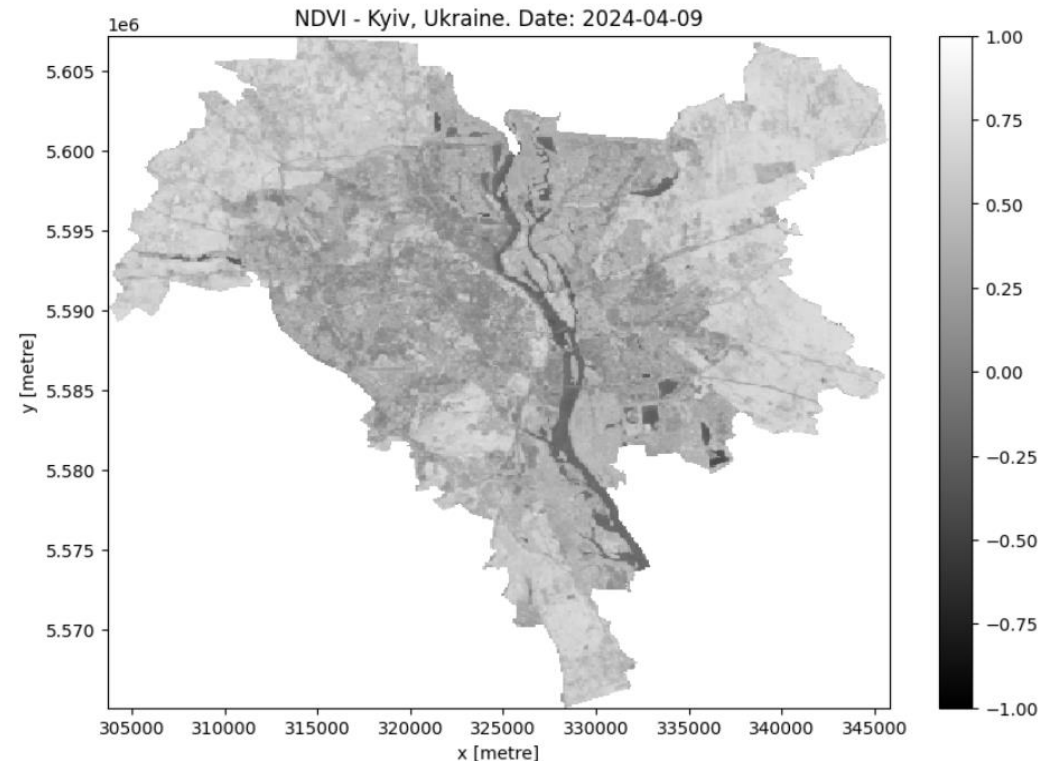
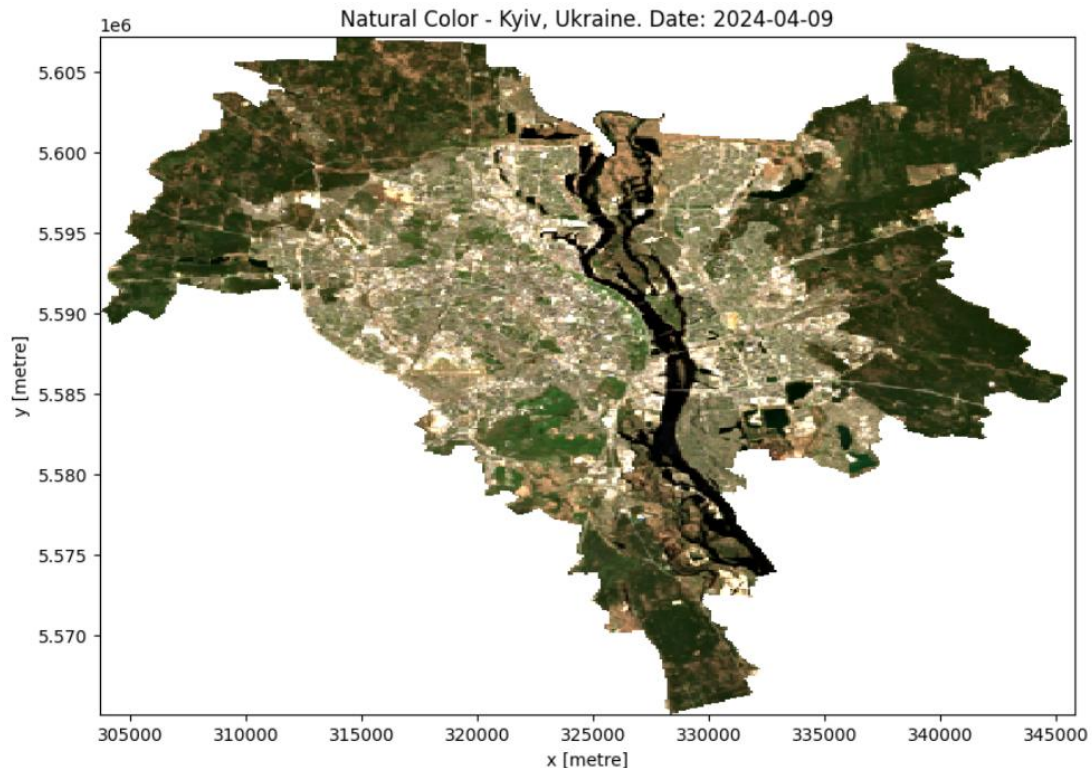
ImportError: cannot import name '_PosixFlavour' from 'pathlib'.

Для усунення цієї проблеми необхідно встановити бібліотеку **pathlib2** і потім, у файлі **urlpath.py**, що знаходиться за шляхом `C:\Users\{user_name}\anaconda3\envs\{env_name}\Lib\site-packages\urlpath.py`, змінити в рядку 31 бібліотеку для імпортування функцій `'_PosixFlavour'` та `PurePath` з **pathlib** на **pathlib2**, як показано на прикладі нижче:

```
urlpath.py
1  #!/usr/bin/env python3
2  # -*- coding: utf-8 -*-
3  """Object-oriented URL from 'urllib.parse' and 'pathlib'
4  """
5  __version__ = '1.2.0'
6  __author__ = __author_email__ = 'brandonschabell@gmail.com'
7  __license__ = 'PSF'
8  __url__ = 'https://github.com/brandonschabell/urlpath'
9  __download_url__ = 'http://pypi.python.org/pypi/urlpath'
10 # http://pypi.python.org/pypi?%3Aaction=list_classifiers
11 __classifiers__ = [
12     'Development Status :: 5 - Production/Stable',
13     'Environment :: Web Environment',
14     'Intended Audience :: Developers',
15     'License :: OSI Approved :: Python Software Foundation License',
16     'Operating System :: OS Independent',
17     'Programming Language :: Python :: 3.4',
18     'Programming Language :: Python :: 3.5',
19     'Programming Language :: Python :: 3.6',
20     'Programming Language :: Python :: 3.7',
21     'Programming Language :: Python :: 3.8',
22     'Programming Language :: Python :: 3.9',
23     'Programming Language :: Python :: 3.10',
24     'Topic :: Internet :: WWW/HTTP',
25     'Topic :: Software Development :: Libraries :: Python Modules',
26 ]
27 __all__ = ('URL',)
28
29 import collections.abc
30 import functools
31 from pathlib2 import _PosixFlavour, PurePath
32 import re
33 import urllib.parse
34
35 try:
36     from unittest.mock import patch
37 except ImportError:
38     from mock import patch
```


3. Використання даних Sentinel-2 для геопросторового аналізу

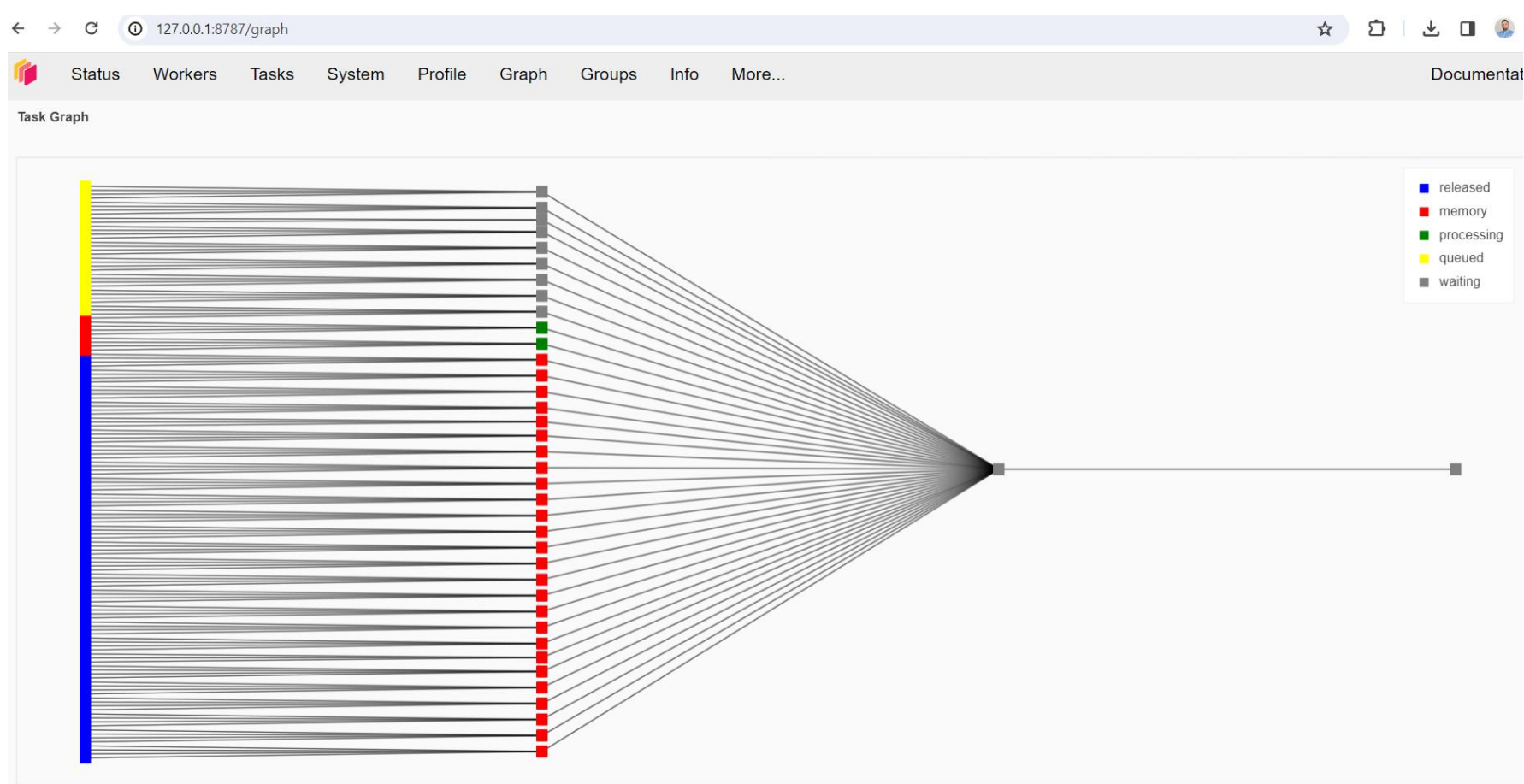
3.1 Даний пункт дозволить закріпити попередньо наданий матеріал і навчить використовувати дані Sentinel-2 для геопросторового аналізу. Для його виконання необхідно виконати всі кроки у файлі **lab_4_1.ipynb**, який знаходиться у папці – **./lab_4**. Завдання буде вважатися виконаним, якщо будуть отримані 2 геозображення, які наведені нижче (для візуалізації NDVI, стар має бути **'RdYlGn'**). Для прикладу, було використано colormap= **'grays'**).



3. Використання даних Sentinel-2 для геопросторового аналізу

Контролювати статус обчислень можна за допомогу Dask інтерфейсу, використовуючи локальний хост :
<http://127.0.0.1:8787/status>

* порт 8787 є за умовчуженням, проте якщо він є зарезервований, необхідно буде використати інший.



- Синій квадрат ("released"): Позначає завдання або ресурси, які були використані та вже не використовуються.
- Червоний квадрат ("memory"): Відображає завдання, дані яких знаходяться в оперативній пам'яті і готові до використання або вже використовуються.
- Зелений квадрат ("processing"): Позначає завдання, які в даний момент активно обробляються.
- Жовтий квадрат ("queued"): Вказує на завдання, які очікують на виконання, тобто вони в черзі і будуть оброблені, як тільки звільняться необхідні ресурси.
- Сірий квадрат ("waiting"): Відображає завдання, які очікують на звільнення ресурсів або завершення інших завдань перед тим, як вони зможуть бути поставлені в чергу або оброблені.

4. Зробити звіт про виконану роботу в електронному вигляді

4.1 Для успішного виконання практичної роботи необхідно підготувати звіт про виконану роботу. У звіті слід детально описати хід виконання кожного пункту завдання та надіслати його у Google Classroom. Також потрібно прикріпити .ipynb файл з виконаними завданнями.