

2024年3月12日,14日,19日,22日

オンライン開催

研究データに関わる  
研究・教育支援人材向け

# Python 活用基礎研修（おわりに）

～研究データ管理基盤GakuNin RDMの紹介を添えて～

附属図書館 研究開発室  
甲斐 尚人・神崎 隼人

主催：附属図書館研究開発室  
共催：産業科学研究所 AI センター  
コアファシリティ機構  
(データ利活用 DX 支援部門・研究支援人材育成部門)  
情報推進部  
サイバーメディアセンター  
協力：技術職員連絡会議  
部局横断型女性技術職員ネットワーク  
産業科学研究所技術室



# Python活用基礎研修 いかがだったでしょうか

4回の研修を通して、研究データ管理に重要な各ライブラリの存在を知っていただきました

	ライブラリ	研究データの処理、クリーニング、可視化に重宝される
入門編 2	NumPy	数値データの効率的な操作と処理に適し、クリーニングプロセスで頻繁に利用
応用編 1	Pandas	データの読み込み、加工、欠損値処理など、データ前処理に関する幅広い機能を提供
応用編 1	Matplotlib (Seaborn)	データの可視化に使われ、クリーニングの過程で発見された問題点の把握に役立つ
応用編 2	Scikit-learn	機械学習モデルを扱う際に用いられ、データの前処理や変換のためのツールも提供



事後アンケートにご協力ください  
忌憚のないご意見をよろしく願っています



研究データ管理基盤

**GakuNin RDM**

をご存知でしょうか

# GakuNin RDMとは??

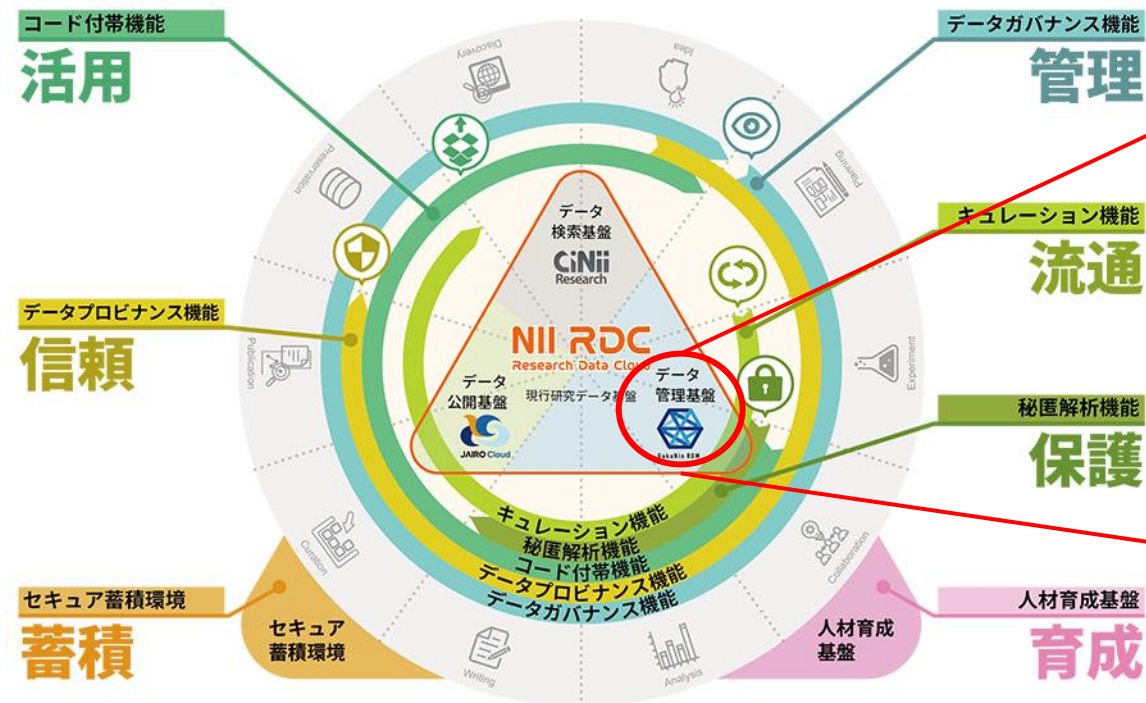
科学技術・イノベーション基本計画

令和3年3月26日  
閣議決定

## (c) 具体的な取組

### ① 信頼性のある研究データの適切な管理・利活用促進のための環境整備

データの管理・利活用のための我が国の中核的なプラットフォームとして  
研究データ基盤システム (NII Research Data Cloud) の普及・広報と



本学構成員であればどなたでもご利用いただけます

<https://support.rdm.nii.ac.jp/>

🔍 サイト内検索

🌐 英語

NII 国立情報学研究所WEBサイト

 **GakuNin RDM** サポートポータル

はじめての方

ご利用中の方

機関管理者の方



ログイン



**研究チームのデータ管理を  
GakuNin RDMに統合しよう。**

GakuNin RDMは、チームを率いる研究者の悩みを解決します。

# コード付帯機能 (Pythonプログラムを実行・共有できる機能)

GakuNin RDM マイプロジェクト 検索 サポート Naoto Kai

テストプロジェクト ファイル Wiki メタデータ 解析 メンバー アドオン 設定 証跡管理

テストプロジェクト 非公開 P 0 ...

メンバー: Naoto Kai  
所属機関: 大阪大学  
作成日時 2023-06-02 02:10 PM | 最終更新日時 2024-03-21 07:26 PM  
カテゴリ: プロジェクト  
説明: プロジェクトに簡単な説明を追加してください  
ライセンス: ライセンスなし

Wiki 🔗

テスト

ファイル 🔗

ストレージプロバイダーをクリックするか、ドラッグ&ドロップしてファイルをアップロードします

🔍 フィルタ i

名前 ^ v	最終更新日時 ^ v
📁 テストプロジェクト	
- 📁 NII Storage	
+ 📁 .binder	
+ 📁 result-u007547e@osaka-u.ac.jp-20240116-5wjzd-osfstorage-xn...	
📄 test.ipynb	2024-01-22 02:25 PM
- 📁 S3 Compatible Storage: testkai2 (osakau)	
📄 スウェーデン研究機関RDM調査報告.pdf	2023-06-02 02:24 PM

引用 ▼

コンポーネント コンポーネントを追加 プロジェクトをリンク

コンポーネントを追加して、プロジェクトを整理します。

タグ

タグを追加してプロジェクトを発見しやすくする

最近の活動 更新 フィルタをクリア

開始日  終了日

ユーザでフィルタ (現在のメンバーから選択)

過去のプロジェクトメンバーも含めて検索する

# Jupyter Labを立ち上げ、処理を実行

The screenshot displays the Jupyter Lab interface. The top menu bar includes File, Edit, View, Run, Kernel, **Tools** (circled in red), Settings, and Help. The left sidebar shows a file browser with a search bar and a list of files, including 'test.ipynb' which is selected. The main area shows a terminal window with the following output:

```
[28]: pip install numpy
Requirement already satisfied: numpy in /srv/conda/envs/notebook/lib/python3.9/site-packages (1.26.3)
Note: you may need to restart the kernel to use updated packages.

[12]: pip install pandas
Collecting pandas
  Downloading pandas-2.2.0-cp39-cp39-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (13.0 MB)
-----
13.0/13.0 MB 71.3 MB/s eta 0:00:00:01:00:01
Requirement already satisfied: python-dateutil>=2.8.2 in /srv/conda/envs/notebook/lib/python3.9/site-packages (from pandas) (2.8.2)
Requirement already satisfied: numpy<2, >=1.22.4 in /srv/conda/envs/notebook/lib/python3.9/site-packages (from pandas) (1.26.3)
Requirement already satisfied: pytz>=2020.1 in /srv/conda/envs/notebook/lib/python3.9/site-packages (from pandas) (2022.7.1)
Collecting tzdata>=2022.7
  Downloading tzdata-2023.4-py2.py3-none-any.whl (346 kB)
-----
346.6/346.6 kB 22.5 MB/s eta 0:00:00
Requirement already satisfied: six>=1.5 in /srv/conda/envs/notebook/lib/python3.9/site-packages (from python-dateutil>=2.8.2->pandas) (1.16.0)
Installing collected packages: tzdata, pandas
Successfully installed pandas-2.2.0 tzdata-2023.4
Note: you may need to restart the kernel to use updated packages.

[14]: pip install pyarrow
Collecting pyarrow
  Downloading pyarrow-15.0.0-cp39-cp39-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (38.4 MB)
-----
38.4/38.4 MB 45.1 MB/s eta 0:00:00:01:00:01
Requirement already satisfied: numpy<2, >=1.16.6 in /srv/conda/envs/notebook/lib/python3.9/site-packages
Installing collected packages: pyarrow
Successfully installed pyarrow-15.0.0
Note: you may need to restart the kernel to use updated packages.

[14]: pip install matplotlib
Collecting matplotlib
  Downloading matplotlib-3.8.2-cp39-cp39-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (11.6 MB)
-----
11.6/11.6 MB 46.9 MB/s eta 0:00:00:01:00:01
Requirement already satisfied: numpy<2, >=1.21 in /srv/conda/envs/notebook/lib/python3.9/site-packages (from matplotlib) (1.26.3)
Requirement already satisfied: packaging>=20.0 in /srv/conda/envs/notebook/lib/python3.9/site-packages (from matplotlib) (23.0)
Collecting contourpy>=1.0.1
  Downloading contourpy-1.2.0-cp39-cp39-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (310 kB)
-----
311.0/311.0 kB 21.2 MB/s eta 0:00:00
Collecting pyparsing>=2.3.1
```

The bottom right corner shows a line chart with the x-axis labeled 'Days' (Day\_1 to Day\_10) and the y-axis labeled 'City' (City\_7, City\_8, City\_9, City\_10). The chart displays multiple colored lines representing different cities over time.



# コード付帯機能 (Pythonプログラムを実行・共有できる機能)



The screenshot shows a project page for 'テストプロジェクト' (Test Project) on GakuNin RDM. The page includes a header with the project name and a blue logo. Below the header, there is a 'Wiki' section with a 'テスト' (Test) entry. The main section is 'ファイル' (Files), which contains a file list. The file list has columns for '名前' (Name) and '最終更新日時' (Last Updated Time). Two files are highlighted with red circles: 'output\_result\_sample\_rainfall.csv' (updated 2024-01-22 02:56 PM) and 'test.ipynb' (updated 2024-01-22 02:25 PM). The file list also shows folders like '.binder' and 'result-u007547e@osaka-u.ac.jp-20240116-5wjzd-osfstorage-xn...'. A search bar with 'フィルタ' (Filter) and an information icon 'i' is visible above the file list.

メンバー: Naoto Kai  
所属機関: 大阪大学  
作成日時 2023-06-02 02:10 PM | 最終更新日時 2024-03-21 07:50 PM  
カテゴリ: プロジェクト  
説明: プロジェクトに簡単な説明を追加してください  
ライセンス: ライセンスなし

Wiki

テスト

ファイル

ストレージプロバイダーをクリックするか、ドラッグ&ドロップしてファイルをアップロードします

Q フィルタ i

名前	最終更新日時
テストプロジェクト	
- NII Storage	
+ .binder	
- result-u007547e@osaka-u.ac.jp-20240116-5wjzd-osfstorage-xn...	
+ .ipynb_checkpoints	
output_result_sample_rainfall.csv	2024-01-22 02:56 PM
+ result-u007547e@osaka-u.ac.jp-20240321-5wjzd-osfstorage-xn...	
test.ipynb	2024-01-22 02:25 PM

Jupyter Notebook (Jupyter Notebook) ファイルや出力結果を GakuNin RDM のストレージに書き出すことが可能

## GakuNin RDM を利用するメリット

例えば...

研究室で GakuNin RDM 上にプロジェクトを立ち上げ、メンバー間のみでファイルを共有

- 先輩から引き継いだ実験の再現
- データクリーニング用のプログラムの共有と再現

等が可能になる

ぜひ

# GakuNin RDM

をご利用いただき、  
フィードバックをお寄せください



Python活用基礎研修、お疲れさまでした

今後の活動にもご期待ください