

はじめよう、研究データ管理

九州大学データ駆動イノベーション推進本部研究データ管理支援部門

九州大学附属図書館図書館DX支援室

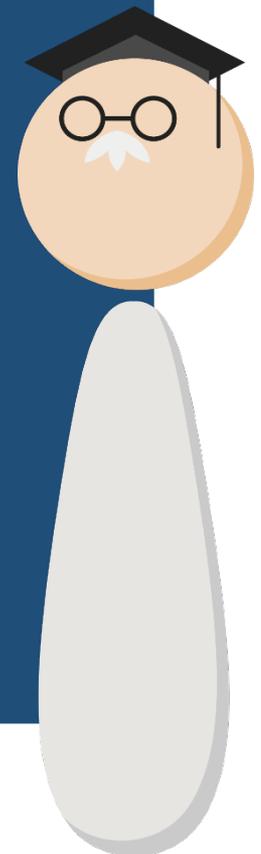
<https://doi.org/10.15017/6796147>

4. 研究データの保存

大切な研究データを長期的に保存できるように、適切なデータの保存方法を知りましょう。

本章では、以下について学びます。

- 4.1 適切な保存の重要性
- 4.2 保存に関する所属機関のポリシー
- 4.3 保存場所
- 4.4 バックアップ
- 4.5 バージョン管理
- 4.6 フォーマット
- 4.7 セキュリティ対策



適切な保存の重要性

- 現在、研究データの多くはデジタルファイル。
- 人的ミスや機械的故障などにより大切な研究データを失うリスクあり。
 - ファイルの意図しない削除・上書き
 - ハードウェアの故障による破損
 - モバイルPCや保存メディアの紛失

研究データを失うことは研究の進展に大きく影響します。
トラブルを避けるため、研究データを適切に保存することが大切です。

所属機関のポリシー

研究データの保存や管理に関するポリシーを定める研究機関が増えて
います。
研究を始める前に、自身が所属する研究機関のポリシーを確認しましょう。

例: 京都大学



『京都大学における公正な研究活動の推進等に関する規程』第7条

- 「教職員等は、適正な保存方法により、一定期間研究データを保存し、必要に応じて当該研究データを開示しなければならない。」

http://www.kyoto-u.ac.jp/uni_int/kitei/reiki_honbun/w002RG00001165.html

九州大学の研究データポリシー・ガイドライン

九州大学では、研究データの管理、公開、保存等に関して以下を定めています。重要なポリシーですから、事前に確認しておきましょう。

• 九州大学研究データ管理・公開ポリシー, 同解説

(令和5年3月16日 教育研究評議会 決定)

より善き知の探究と創造・展開の拠点として人類と社会に貢献することを使命とする九州大学において、新たな知の源泉となる研究データを効率的に管理し、その利活用を促進することで新たな知の創出を目指すという理念の下、研究データの管理及び公開の原則を定める

- **ポリシー** https://rds.dx.kyushu-u.ac.jp/wp-content/uploads/2023/04/KU_DataPolicy.pdf
- **解説** https://rds.dx.kyushu-u.ac.jp/wp-content/uploads/2023/04/KU_DataPolicy_commentary.pdf

• 研究データの保存等に関するガイドライン

(実施:平成27年8月18日)

「研究データ等のうち、実験ノート、数値データ、画像等、「資料」の保存期間は、原則として、当該論文等の発表後10年間とする」と定める

<https://www.kyushu-u.ac.jp/f/1461/guideline.pdf>

データの保存場所(1/2)

容量や費用、簡便さに加え、信頼性や安全性など各保存場所の特徴を考慮し、保存先を選択しましょう。

代表的な4つの保存場所

1. デスクトップ・モバイルPC

簡便な保存先ではあるが、ハードウェアが冗長化されておらず、障害によりファイルを失う可能性大。研究資料のマスターコピーを保存する場所としては不適切。

2. 外部記憶デバイス

CDやDVDはサイズが十分ではなく、書き込みエラーへの対処や確認が必須。USBハードディスクやUSBメモリ、SDカードなどは、安価で大容量化してきたが、PC・ラップトップと同様の問題あり。ポータビリティに優れる反面、紛失や盗難などのセキュリティリスク大。

データの保存場所(2/2)

3. クラウドストレージ (例:OneDrive, Dropbox)

インターネット上のディスクスペースを利用できるサービス。一定容量までは無料で利用できるものもあるが、機関のセキュリティポリシーやガイドラインに適合するものを選択すべき。

*オンラインストレージサービスの比較:

https://en.wikipedia.org/wiki/Comparison_of_online_backup_services

4. ネットワークドライブ (例:研究室のNAS)

LANなどのネットワーク経由で接続できる他のコンピュータのディスクスペース。安定的に運用される所属機関のシステム管理部門が提供するものを利用するのが理想。

外部のクラウドストレージを利用する際は、機関のセキュリティポリシーに合致するかを事前に確認しましょう。所属する研究機関が提供するネットワークドライブ、クラウドストレージがあれば、安心して研究データを保存できます。クラウドストレージやネットワークドライブ等を利用して複数の環境で作業する際には、確実にファイルを同期させるようにしましょう。

ストレージ

九州大学では、以下のストレージを利用できます。

• QRDM

- 研究データ管理用ストレージシステムとして2022年度に導入、2023年サービス開始
- Nextcloudによるデータの効率的な管理
- GakuNin RDMと連携 (次スライド参照)

• Proself, OneDrive for Business

- 情報統括本部が提供
- OneDrive for Businessは、GakuNin RDMと連携可能 (次スライド参照)

クラウドサービスを導入する場合は、「九州大学クラウドサービス利用ガイドライン」によるチェックを行い、サービス利用の可否、利用するサービスの選定などを検討してください。検討の結果、クラウドサービスを利用することを決定した場合は、結果報告書を作成し、所定の添付資料とともに情報統括本部へ提出してください。

<https://cloud.iii.kyushu-u.ac.jp/>

GakuNin RDM

九州大学の構成員は、GakuNin RDMを利用できます。

- **GakuNin RDM**とは
 - <https://rdm.nii.ac.jp>
 - 国立情報学研究所が提供する研究データ管理基盤
- GakuNin RDMでできること
 - バージョン管理
 - アクセスコントロール
 - 研究証跡の記録
 - QRDM、OneDrive for Businessと連携
- ログイン方法
 - GakuNin RDMにアクセス > 所属している機関として「九州大学」を選択 > SSO-KIDでログイン
 - お問い合わせ：研究データ管理支援部門
https://rds.dx.kyushu-u.ac.jp/contact_us

研究データのバックアップ(1/3)

不慮のトラブルで研究データを失った際にも、バックアップデータがあれば、トラブルの影響を最小限に抑えることができます。
定期的にバックアップを取りましょう。

- バックアップの場所
 - ネットワークドライブ、クラウドストレージなどの選択肢から、研究内容に応じて適切なものを選択
 - 情報セキュリティや災害リスクも考慮
 - 絶対に漏洩してはいけないデータは、オンライン上ではなく、ローカルのコンピュータやストレージに保管
- 複数のバックアップを取る
- 機密性のレベル等について各機関が提供する情報を確認する

研究データのバックアップ(2/3)

• バックアップの方法

- 手動ではなく、バックアップソフトウェアなどを利用し、定期的なバックアップを取得する
 - 定期的にバックアップされていることを確認。
 - バックアップからリストアできることを確認。
- ディスク容量やバックアップに必要な時間を勘案し、フルバックアップ、増分・差分バックアップから適切な方法を選択する
 - ファイルの更新頻度などから、適切なバックアップ頻度を設定。
例) 日次:差分バックアップ、月次:フルバックアップ

研究データのバックアップ(3/3)

- 災害等を考慮し、複数の場所にバックアップを保管しましょう。
- 最低でも**3つのコピー**があることが推奨されています。
 - 「**オリジナル(original)**」: 最初のコピー、作業用のデータセットと関連ファイル。通常、研究者が主に使用するコンピュータに保管。
 - 「**ニア(near)**」: オリジナルの近くになければならないが、物理的な場所が同じでないことが理想的。このコピーは自動バックアップソフトか手動のどちらかで更新。多くの場合、ニア・コピーは外部ハードドライブや研究者の属する施設内の共有ファイルサーバに保管。
 - 「**ファー(far)**」: オリジナル・コピーからもニア・コピーからも物理的な場所が遠い。災害の脅威が異なる場所に置くのが理想。

バージョン管理

- 多くの場合、研究者は独自の方法を利用し、データセットを始めとするファイルのさまざまなバージョンを識別。
- 適切に設計されたバージョン管理システムを利用すれば、ファイルに加えられた変更をバージョン番号、タイムスタンプ、変更箇所の説明とともに記録できる。変更は容易に比較することができ、必要に応じて復元することも可能。
- バージョン管理システムの一般的な例…Git

ファイルフォーマット

論文を執筆するまでの短期的な視点ではなく、長期的に研究データが利用できる状態を確保するためには、フォーマットによる特徴の違いを考慮し、適切なフォーマットを選択する必要があります。

- 推奨されるファイルフォーマット
 - 特定の商用ソフトウェアから独立したフォーマット
 - 仕様が国際標準化されているフォーマット
 - 広く普及しているフォーマット
- フォーマット変換のリスク
 - フォーマットを変換することにより、ファイル内部に記録されたメタデータや画像の品質、文章構造などの情報を失う可能性大。
- テキストとバイナリ
 - テキストファイルは環境に依存せず開くことができ、人間可読性高。
 - バイナリファイルはよりリッチな情報を包埋しているが、人間可読性低。

セキュリティ対策

セキュリティポリシーに沿った研究活動や環境を整備することで、トラブルを回避し、総合的には円滑な研究活動に繋がります。

- 物理的なセキュリティ対策
 - 離席する際のコンピュータをロック
 - 部屋の施錠
 - モバイルPCを放置しない
 - USBメモリ等を紛失しないための対処
- PC上でのセキュリティ対策
 - アンチウイルスソフトウェア／ファイアウォールの導入
 - ソフトウェアやOSのアップデート
 - 適切な強度のパスワードや認証方法の利用
 - ファイルやハードディスク、USBメモリ等の暗号化

九州大学の皆様へ

- 本章で示した研究データの保存に関するルールは、研究が始まる前にあらかじめ決めておいたほうがよいでしょう。
 - 研究室や共同研究など複数人で研究を進める際には、研究プロジェクトごとにルールを決めましょう。
 - 研究の遂行中に、変更が生じた場合も、その都度、相談するようにしましょう。
 - 個人情報などが含まれる研究データを取り扱う場合は、研究データの保存先は慎重に選択しましょう。
- ご不明な点があれば、以下に相談してください。
研究データ管理支援部門
https://rds.dx.kyushu-u.ac.jp/contact_us
rds_help@dx.kyushu-u.ac.jp

