

1-3.

研究データ、研究データ管理の定義



研究データとは

研究データとは研究の過程で、あるいは研究の結果として収集・生成される情報である。仮説を検証するために使用されたり、結論を導くための根拠となる。また、後に研究結果が正しいかどうか確認するために使用されることもある。

電子的なデータのみ研究データと定義することも、紙ベースの情報も研究データに含める場合もある。

研究データのとり形式はテキストや3Dモデルなど様々であり、記録されるものも研究ノートや写真、音声テープなど多岐にわたる。

メモ

研究データを構成するもの

文書ファイル、スプレッドシート

研究ノート、フィールドノート（野外調査の記録）、日誌

アンケート、転写物、コードブック

音声テープ、ビデオテープ

写真、フィルム

試験・検査への反応、応答

スライド、人工物、標本、サンプル

研究の過程で獲得され、生成されたデジタル資料のコレクション

統計的なデータファイル

データベースのコンテンツ（ビデオ、音声、テキスト、画像）

モデル、アルゴリズム、スクリプト

アプリケーションソフト

方法論、ワークフロー

標準的な操作手順

2-1も参照

メモ

研究データの具体例

観測データ (Observational)

自然現象等を観測し取得されたデータ。例：天体観測、気象観測、社会調査

実験データ (Experimental)

実験によって得られたデータ。高価になることもある。

例：DNAシーケンス、クロマトグラム、微量検定法により得られたデータ

シミュレーション (Simulation)

モデルやコードを用いて生成されたデータ。得られたデータに加え、使用されたモデルやコード、パラメータセットも重要。例：気象モデル、経済モデル

メモ

研究データの具体例

派生データ・編集されたデータ (derived/compiled)

取得したデータを加工、あるいは結合することによって得られたデータ。

例：データマイニング、テキストマイニング、3Dデータ

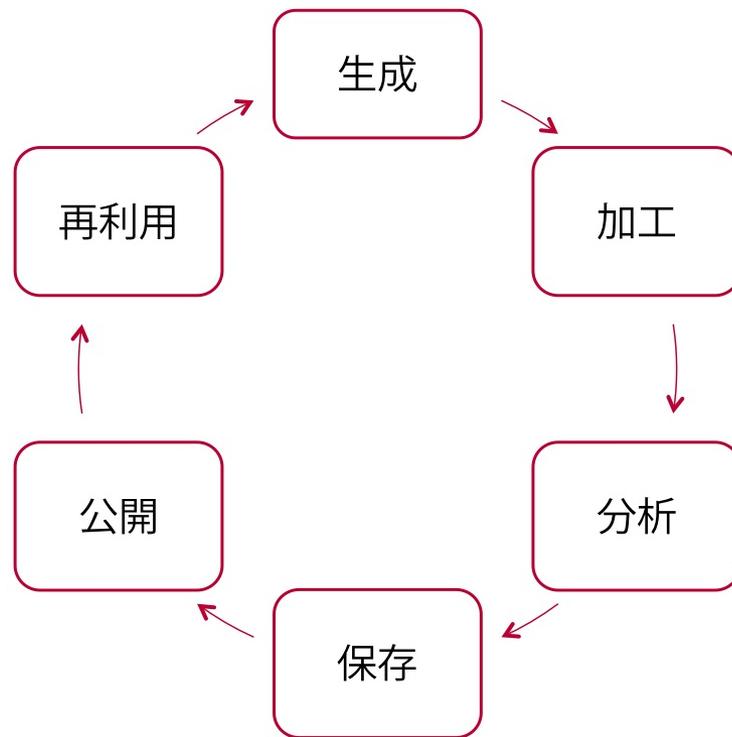
参照/標準データ (Reference or canonical)

研究の過程で参照されるデータ、データセット。あるいは標準的なデータのコレクション。整理・共有されているものも多い。

例：遺伝子配列データバンク、結晶データベース、歴史画像アーカイブ

メモ

研究データのライフサイクル



- The UK Data Archive (<http://www.data-archive.ac.uk/create-manage/life-cycle>)
- RDM Support - basic training course for information specialists(<http://dx.doi.org/10.6084/m9.figshare.1285313>)より

メモ

研究と研究データ

研究データはしばしば、その研究データが得られた研究プロジェクトよりも寿命が長い。研究者は資金援助が終わった後もデータに関する作業を続け、後に続く研究ではそのデータが分析されたり、新たに追加され、さらには他の研究者によって再利用されることもありうる。

組織化、文書化、保存・共有がしっかりとされたデータは科学研究を推進することや、学習及びイノベーションの機会を増やすことにおいて計り知れない価値がある。

<https://www.ukdataservice.ac.uk/manage-data/lifecycle> より

メモ

研究データ管理 (Research data management) とは

「研究データ管理」とは、一般的にある研究プロジェクトにおいて使用された、あるいは生成された情報を、どのように組織化、構造化、保管、管理していくのかを指す言葉。

<http://researchdata.ox.ac.uk/home/introduction-to-rdm/> より

メモ

研究データ管理 (Research data management) とは

研究データ管理には次のようなことが含まれる

どのようにデータを取り扱っていくかの計画 (Data Management Plan) 策定

研究プロジェクトが行われている間の、日々の情報の取り扱い方

長期的にデータをどのように取り扱うか (研究プロジェクトが終わった後はそのデータをどうするのか)

1-4~1-6も参照

<http://researchdata.ox.ac.uk/home/introduction-to-rdm/> より

メモ

研究データ管理のメリット

研究効率

研究効率の上昇

長期的に時間や資源を節約

データセキュリティの強化とデータ消失リスクの最小化

データを他の研究者が再利用することにより、労力の重複を避ける

ある時点のデータを他の時点と比較できるデータを提供することにより、データの変化の分析を容易にする

ポリシー

資金提供元の要求を満たす

産業界や商業界の慣行に従う

研究倫理

研究の再現性の確立

研究データや研究記録の正確性、完全性、真正性、信頼性の保証

メモ

研究データ管理と大学図書館

すでに海外の大学ではデータ管理計画作成等の研究データ管理業務に関する支援サービスを図書館や他の関連組織が共同で行っているところもある

4-2,4-3も参照

例 1 バージニア大学 <http://data.library.virginia.edu/data-management/>

例 2 エディンバラ大学 http://www.ed.ac.uk/files/atoms/files/rds_booklet_may2016.pdf

例 3 ケンブリッジ大学 <http://www.data.cam.ac.uk/support>

例 4 アムステルダム大学 <http://rdm.uva.nl/en>

メモ