

研究データマネジメント(RDM)用の ルーブリック -- Basic 版 --

*家森俊彦¹, 梶田将司¹, Janice Smith², Jacques Raynauld²
能勢正仁³, 青木学聡⁴, 原正一郎¹, 宮野公樹¹

Rubric for improving research data management (RDM) skills -- Basic version --

T. Iyemori¹, S. Kajita¹, J. Smith², J. Raynauld²
M. Nose³, T. Aoki³, S. Hara¹ and N. Miyano¹

¹Kyoto University, ²Karuta Project,
³Nagoya City University, ³Nagoya University



内容

1. Basic版ルーブリックの役割と作成
2. ルーブリックの構成
3. 研究の進捗とルーブリックの項目

Basicルーブリックの役割と作成

1. よりよい研究のための準備
 - ・ どのようなデータが必要か?
 - ・ データの取得と処理方法は?
 - ・ 必要な研究費の額は?
 - ・ DMP (Data Management Plans) 提出の必要
2. 研究進捗状況のチェック
 - ・ データの取得と保管状況
 - ・ データ処理・解析状況
 - ・ 今後の必要にそなえた文書化
3. 有意な解析結果を得る
 - ・ 必要なソフトウェアはあるか?
 - ・ 品質の良いデータが用意されているか?
(例: ノイズ除去)
4. オープンサイエンス(データ)時代への対応

RDM用ルーブリックの作成

目標: 研究の進行あるいはライフサイクルの各段階で RDMに関する知識およびスキルの学習あるいは自己評価ができる仕組みの作成

分野: 様々な分野向けのルーブリックを作成。

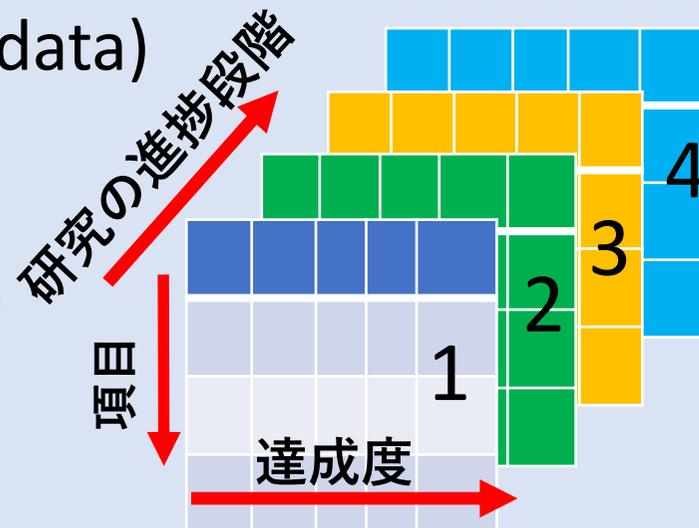
それらを基に、できる限り一般化する。 → **Basic rubric**

対象: 主に学生・若手研究者

特徴: 研究の進捗段階に対応した4つのルーブリックに分割、すなわち、

1. データに関する計画(Planning for data)
2. データの処理と整理(Organizing data)
3. データの解析(Analyzing data)
4. データの共有または公開
(Sharing and publishing data)

ルーブリックの各項目毎に4段階の達成度を設定



研究の進捗段階

1. 計画段階
↓
Data Management Plan (DMP)
2. 研究開始
↓
Data acquisition
↓
Preparation for analysis
3. 解析
↓
Data analysis
4. 解析結果とその検討
5. 論文作成と発表
↓
学術誌への投稿

対応するルーブリック



PLANNING FOR DATA

データに関する計画



ORGANIZING DATA

データの取得と処理・保管



ANALYZING DATA

データ解析



SHARING OR
PUBLISHING DATA

データの公開と再利用

現在までに作成した分野別ルーブリック

(特徴)

- ・ 地球惑星科学分野 – 主に観測データを念頭に –
- ・ 物質科学(材料)分野 – 主に室内実験データを念頭に –
- ・ 人文・社会科学(地域研究)分野 – 主に収集資料 –
- ・ 学際研究用 – 研究グループの形成方法に関するルーブリックを追加 –

(作成予定)

- ・ 生物(植物)学分野

かなりの部分は共通している → **Basic rubric**

データに関する計画 PLANNING FOR DATA

各項目の達成度



データに関する計画 PLANNING FOR DATA	初期 Beginning	発展 Developing	向上 Enhancing	完了 Completed
取得するデータ Data to Be Obtained	取得するデータとそれに必要な素材/対象が決まりました。	取得するデータの種類と、各研究目的との関係が特定されています。	データの解像度、品質、最小データ量、およびデータ取得に必要な対象を特定しました。	データの種類、解像度、品質、必要最小量の見積もりに自信があり、必要なデータを確実に取得できると確信しています。
データの取得方法 Methodology for Obtaining Data	データを取得するために必要な方法を決めました。	データの取得と保存に必要な機器、データ転送計画、およびストレージデバイスを決めました。	機器、データ転送の仕組み、および記憶装置の準備または実装を始めました。	機器、データ転送の仕組み、およびストレージデバイスの準備または実装が完了しました。
データ収集のための材料、機器、およびソフトウェア Materials, Equipment, and Software for Data Acquisition	資材、設備、ソフトウェアの購入あるいは作成の計画を決めました。	データを収集/生成する場所が決まり、その場所での作業の準備が整いました。	材料、機器、あるいはソフトウェアに関連する国内および国際的に必要な法的考慮事項が特定され、遵守されています。	データの収集や生成に必要な材料、機器、ソフトウェアを入手する手順が文書化され、完了しました。
研究グループ構成員 Team Members	データの取得と処理、保管、配布等を担当する研究グループの構成員が決まりました。	各構成員のスキルと興味に応じて、適切な役割が割り当てられています。	研究グループ構成員間の情報交換、相談、協力を促進するために、適切な研究環境が準備されています。	各構成員は役割分担と自分の役割を十分に理解しており、研究を実施し成果を公表するための十分な準備ができています。
データ取得のための資金 Funding for Data Acquisition	データ収集に必要な資金が見積もられ、利用可能な資金源が特定されました。	資金申請には、構成員の活動経費、機器や材料の購入、データ収集用のソフトウェア、法的に必要な経費などが含まれます。	資金のリクエストに関連する資金提供組織に提出しました。	データ取得に必要なすべての費用に十分な資金を獲得しました。
データの再利用 Data Re-Use	研究データを再利用するために管理方法の変更が必要となるか検討しました。	他の研究者がデータを再利用できるかどうかを決定しました。	研究機関で定められた期間の間は研究グループの構成員がデータを独占的に使用する権利に関して決定がなされました。	データの管理に必要な変更を含め、データの再利用の可能性に関する計画の作成が完了し、研究グループ構成員全員が承認しました。
データの公開・非公開 Open Versus Closed Data	研究成果を公表する可能性のある学術雑誌を特定し、そこでのデータ開示の条件を理解しています。パブリックリポジリでのデータの共有を検討しています。	どのデータ（および資料）を公開および共有できるかできないかを決めました。関連する国際法、現地法、およびデータ所有権についても考慮しています。	データを公開および共有するためのライセンスとリポジトリの選択、およびデータ（および資料）の潜在的なユーザーについて検討しました。	データ（およびマテリアル）を公開および共有するためのライセンスとリポジトリの選択結果が文書化され、研究グループ構成員全員によって同意されています。
データの公開と共有の計画 Planning for Publishing and Sharing Data	データ（および資料）の開示と管理をするにあたり、資金提供機関および研究機関が要求している条件を理解しています。	データ（および資料）の開示に必要な資金は見積もられています。データ公開の責任を研究者の所属機関が負うのか、外部委託するのかを決めました。	もし必要なら、公開および共有プロセスをアウトソーシングするための適切なリポジトリが特定され、費用を継続的にどこが負担するかを決めました。	データ（および資料）を公開および共有するための研究データ管理（RDM）計画が作成され、文書化され、研究グループ全構成員と研究者の所属機関が同意しています。

データの処理と整理 ORGANIZING DATA

データの処理と整理 ORGANIZING DATA	初期 Beginning	発展 Developing	向上 Enhancing	完了 Completed
メタデータ Metadata	研究目的達成に必要な研究データ（および資料）のメタデータのモデルの検討を始めた。	メタデータの項目、データ型、記述ルールを決めました。	メタデータを記述するための機械可読形式を選択しました。	メタデータは国際標準の形式に変換され、メタデータを定義するプロセスも文書化されています。
データ収集 Data Collection	観測と測定項目の計画を決定しました。	計画した観測と測定の詳細は、他の研究者が理解しやすい言語で文書化されています。	観測と測定が実行され、結果として得られたデータが機械可読形式に変換されました。	データを安全なサーバーに定期的にアップロードする計画が立てられ、実行されています。
研究資料の入手 Materials Acquisition	必要に応じて、研究資料を入手するための手順が決定され、文書化されています。	必要に応じて、既存の研究資料を補足する追加資料を作成する計画が作成され、文書化されています。	必要に応じて、研究資料を安定した安全な施設に定期的に保管する計画が決定されています。	必要な研究資料が取得および/または作成および保管されている。
デジタル化 Digitization	必要に応じて、データ（および資料）をデジタル化する手順が作成され、文書化されています。	データ（および資料）のデジタル化に関する仕様のリストが作成され、文書化されています。	仕様リストのすべての項目は、他の研究者が理解できる言語で書かれています。	デジタル化されたデータ（および資料）を安定した安全なサーバーに定期的にアップロードする計画が策定されています。
データストレージ Data Storage	データ（およびデジタル化された資料）を保存するために必要な容量を見積もりました。	必要なストレージ容量を確保するための手配を行ないました。	データ（およびデジタル化された資料）が予想以上に増加した場合でも、十分なストレージ容量が利用可能です。	研究組織の構成員は、データ（およびデジタル化された資料）が十分な期間にわたって安全に保管されると確信しています。
物理的資料の保管 Storage of Physical Materials	必要に応じて、物理的（デジタル化された数値情報ではない）研究資料を保管するための設備、保管に必要な環境（温度、湿度など）、および保管にかかる費用を見積もりました。	物理的研究資料を保管するために必要な機器と施設を確保するための手配を行いました。	物理的研究資料の保管に必要な容量は見積もられており、計画以上に増加した場合でも対応可能です。	研究グループ構成員は、物理的研究資料が十分な長期間にわたって安全に保管されると確信しています。
ファイル形式 File Format	この分野で一般的に使用されている国際的に受け入れられたファイル形式を利用することが考慮されています。	データ解析を効率的に行うために、わかりやすいファイル形式を選択しました。	選択したデータ形式が文書化されています。	データは国際的に流通している形式でフォーマットされています。
バージョン管理 Version Control	複雑なデータを長期にわたって処理するためのバージョン管理計画を作りました。	データの最新バージョンを以前のバージョンから区別するための記述子を決めました。	以前のバージョンのデータを保存するためのアーカイブシステムを準備しました。	研究組織構成員は、既存のすべてのバージョンのデータが、それらにアクセスしたい他の研究者に利用可能になっていると確信しています。
バックアップ Backup	データをバックアップするためのさまざまなオプションが検討され、利用可能な資金あるいは組織のリソースが特定されました。	データを自動的にバックアップするためのリソースが利用できない場合は、手動バックアップの手順が用意されています。	起こり得る緊急事態に備えて、データを複数の場所またはリポジトリにバックアップするための手配が行われています。	研究者は、事故による損失の可能性に関係なく、研究データが利用可能であると確信しています。

データの解析 ANALYZING DATA

データの解析 ANALYZING DATA	初期 Beginning	発展 Developing	向上 Enhancing	完了 Completed
解析用ソフトウェア Software for Analysis	データ解析のために、必要な機能を備えた適切なソフトウェアを特定し、入手しました。	利用可能な既存のソフトウェアに研究に必要な機能の一部が欠けている場合には、それを補うために「自家製」ソフトウェアを開発しました。	利用予定のソフトウェアの解説書を入手するとともに、「自家製」ソフトウェアの説明文書を作成しました。	研究グループの構成員は、入手あるいは開発したソフトウェアが研究データの適切な解析を提供すると確信しています。
データの処理と解析 Data Processing and Analysis	すべてのデータ（および資料）はデジタル化の計画に従ってデジタル化されており、安全なサーバーに保管されています。	有意義な結果が得られるように、データは定性的および定量的に解析されています。	解析されたデータは、専門分野で一般的に使用される形式でフォーマットされています。	データの処理・解析方法やフォーマットを説明した文書を作成しました。
中間データ Intermediate Data	研究過程で生成された中間データが保存および保管に値するかどうかについて検討しました。	保存に値する場合は、中間データのフォーマットと保存に関する計画を立てます。	保持に値する中間データが取り出され、保存されます。	中間データに関する説明資料を作成しました。
データクリーニング Data Cleaning	データは視覚的にスキャンされ、ノイズ（エラー、スパイク、階段状のジャンプ等）と考えられる部分を特定・検出し、分類・整理しました。	（ノイズを処理し訂正された解析用のデータファイル作成には）元のデータファイルの複製が使われ、元のデータファイルは保持されます。	複製されたファイル内のすべてのノイズは分類され、削除あるいは修正されました。ノイズの分類方法が文書化されています。	ノイズ訂正済みのデータファイルを再検査し、間違っていると誤って判断されたデータが特定された場合は正しい元のデータに置き換えました。

データの共有または公開 SHARING OR PUBLISHING DATA

データの共有または公開 SHARING OR PUBLISHING DATA	初期 Beginning	発展 Developing	向上 Enhancing	完了 Completed
ライセンス Licensing	他の研究者によるデータの利用に関するルールを定めました。	国際的に認められた適切なライセンスを選択しました。。	選択したライセンスの使用要件は、研究グループ全構成員によって理解され、同意されています。	選択したライセンスの表示が公開データに添付されました。
メタデータスキーマ Metadata Schema	研究に必要なメタデータはどのようなものかを認識しており、その作成を容易にするために利用可能なソフトウェアを特定しました。	この分野で国際的に使用されているメタデータスキーマを選択しました。。	必要なメタデータとその重要性、使用方法、パラメータは、選択したメタデータスキーマに従って記述しました。	メタデータとそのスキーマが完全であり、他の研究者がデータ（およびマテリアル）を利用するのを支援する準備ができていますと確信しています。
データセットへのマテリアルの収集 Collection of Materials into Datasets	研究資料をデータセットに編成するためのメタデータモデルが選択され、適応されています。	研究資料を説明するためのメタデータに必要な項目、データ型、ルールが定義され、文書化されています。	研究資料のメタデータ記述のルールは、データが保管されている機関のルールに可能な限り準拠します。正確な説明を保証するために URI (Uniform Resource Identifier) が使用される場合があります。	研究資料のメタデータ記述が、その分野で受け入れられているメタデータ形式と互換性がない場合、データセットの発見、アクセス、使用を効率化するために変換されます。
ドキュメンテーション Documentation	データ（および資料）の公開と共有に必要な文書の完全なリストが特定されました。	ドキュメントには、メタデータの構造を定義するスキーマが含まれています。	ドキュメントには、重要性、値、ファイル名、使用規則、物理量、データを取得するために使用した機器が含まれます。	必要なドキュメントはすべて作成され、データ（およびマテリアル）の公開と共有に伴うドキュメントファイルに追加されています。
データの完全性、有効性、信頼性、合法性、倫理 Data Completeness, Validity, Reliability, Legality, and Ethics	必要に応じて、研究におけるすべての実験および/またはヒト被験者の使用について IRB (治験審査委員会) の承認が得られており、研究に対する倫理的アプローチの必要性が満たされています。	この分野で認められているデータ収集および分析方法を使用することで、データ収集および分析の完全性（または誠実さ）、有効性、信頼性、合法性および倫理性が保証されています。	データ収集と分析の完全性、有効性、信頼性、合法性、倫理性の証拠が文書化されています。	データ収集と分析の完全性、有効性、信頼性、合法性、倫理性を証明するために、第三者によるレビューが実施されました。
研究資料の保存、返却、廃棄 Preserving, Returning, or Disposing of Research Materials	研究資料の保存、返却、または処分を選択肢は、そのために利用可能な資金も含めて検討しました。	研究資料の保存、返却、あるいは処分について決定し、これらの目的のための資金提供またはその他の取り決めが得られました。	研究グループ全構成員と研究者の所属機関の代表者は、研究資料の保存、返却、あるいは廃棄に関する決定と手順を理解し、受け入れています。	研究資料の保存、返却、処分が完了しました。
データの公開と共有 Publishing and Sharing Data	データ（およびデジタル化された資料）を公開および共有するための RDM 計画の検討を行いました。Web 上のオブジェクトにアクセスするためのデジタルオブジェクト識別子 (DOI) が取得されました。	データの所有権、ネットワークセキュリティ、バックアップシステムに関して、利用する可能性のあるデータリポジトリのポリシー、手順、コストを検討し、公開スケジュールを決定しました。	データ（およびデジタル化された資料）を公開および共有するために、適切なリポジトリを選択しました。資金は確保され、リポジトリとの調整も完了しました。	メタデータおよび文書を含むデータ（およびデジタル化された資料）が公開され、アクセスに関する以前の決定に従って他の研究者が利用できることを確認しました。。

データに関する計画 PLANNING FOR DATA

データに関する計画 PLANNING FOR DATA	初期 Beginning	発展 Developing	向上 Enhancing	完了 Completed
取得するデータ Data to Be Obtained	取得するデータとそれに必要な素材/対象が決まりました。	取得するデータの種類と、各研究目的との関係が特定されています。	データの解像度、品質、最小データ量、およびデータ取得に必要な対象を特定しました。	データの種類、解像度、品質、必要最小量の見積もりに自信があり、必要なデータを確実に取得できると確信しています。
データの取得方法 Methodology for Obtaining Data	データを取得するために必要な方法を決めました。	データの取得と保存に必要な機器、データ転送計画、およびストレージデバイスを決めました。	機器、データ転送の仕組み、および記憶装置の準備または実装を始めました。	機器、データ転送の仕組み、およびストレージデバイスの準備または実装が完了しました。
データ収集のための材料、機器、およびソフトウェア Materials, Equipment, and Software for Data Acquisition	資材、設備、ソフトウェアの購入あるいは作成の計画を決めました。	データを収集/生成する場所が決まり、その場所での作業の準備が整いました。	材料、機器、あるいはソフトウェアに関連する国内および国際的に必要な法的考慮事項が特定され、遵守されています。	データの収集や生成に必要な材料、機器、ソフトウェアを入手する手順が文書化され、完了しました。
研究グループ構成員 Team Members	データの取得と処理、保管、配布等を担当する研究グループの構成員が決まりました。	各構成員のスキルと興味に応じて、適切な役割が割り当てられています。	研究グループ構成員間の情報交換、相談、協力を促進するために、適切な研究環境が準備されています。	各構成員は役割分担と自分の役割を十分に理解しており、研究を実施し成果を公表するための十分な準備ができています。
データ取得のための資金 Funding for Data Acquisition	データ収集に必要な資金が見積もられ、利用可能な資金源が特定されました。	資金申請には、構成員の活動経費、機器や材料の購入、データ収集用のソフトウェア、法的に必要な経費などが含まれます。	資金のリクエストに関連する資金提供組織に提出しました。	データ取得に必要なすべての費用に十分な資金を獲得しました。
データの再利用 Data Re-Use	研究データを再利用するために管理方法の変更が必要となるか検討しました。	他の研究者がデータを再利用できるかどうかを決定しました。	研究機関で定められた期間の間は研究グループの構成員がデータを独占的に使用する権利に関して決定がなされました。	データの管理に必要な変更を含め、データの再利用の可能性に関する計画の作成が完了し、研究グループ構成員全員が承認しました。
データの公開・非公開 Open Versus Closed Data	研究成果を公表する可能性のある学術雑誌を特定し、そこでのデータ開示の条件を理解しています。パブリックリポジトリでのデータの共有を検討しています。	どのデータ（および資料）を公開および共有できるかできないかを決めました。関連する国際法、現地法、およびデータ所有権についても考慮しています。	データを公開および共有するためのライセンスとリポジトリの選択、およびデータ（および資料）の潜在的なユーザーについて検討しました。	データ（およびマテリアル）を公開および共有するためのライセンスとリポジトリの選択結果が文書化され、研究グループ構成員全員によって同意されています。
データの公開と共有の計画 Planning for Publishing and Sharing Data	データ（および資料）の開示と管理をするにあたり、資金提供機関および研究機関が要求している条件を理解しています。	データ（および資料）の開示に必要な資金は見積もられています。データ公開の責任を研究者の所属機関が負うのか、外部委託するのかを決めました。	もし必要なら、公開および共有プロセスをアウトソーシングするための適切なリポジトリが特定され、費用を継続的にどこが負担するかを決めました。	データ（および資料）を公開および共有するための研究データ管理（RDM）計画が作成され、文書化され、研究グループ全構成員と研究者の所属機関が同意しています。

(前のスライドで赤枠で囲った部分の説明)

データ取得のための資金

- ・ 観測装置、実験装置
- ・ 資料の購入
- ・ 旅費
- ・ データ処理装置(WS, HDD, etc.)
- ・ 人件費(謝金 etc.)

データの公開と共有の計画

- ・ 公開に必要な情報(許諾、ライセンス、倫理)をあらかじめ確認

データの処理と整理 ORGANIZING DATA

データの処理と整理 ORGANIZING DATA	初期 Beginning	発展 Developing	向上 Enhancing	完了 Completed
メタデータ Metadata	研究目的達成に必要な研究データ（および資料）のメタデータのモデルの検討を始めた。	メタデータの項目、データ型、記述ルールを決めました。	メタデータを記述するための機械可読形式を選択しました。	メタデータは国際標準の形式に変換され、メタデータを定義するプロセスも文書化されています。
データ収集 Data Collection	観測と測定項目の計画を決定しました。	計画した観測と測定の詳細は、他の研究者が理解しやすい言語で文書化されています。	観測と測定が実行され、結果として得られたデータが機械可読形式に変換されました。	データを安全なサーバーに定期的にアップロードする計画が立てられ、実行されています。
研究資料の入手 Materials Acquisition	必要に応じて、研究資料を入手するための手順が決定され、文書化されています。	必要に応じて、既存の研究資料を補足する追加資料を作成する計画が作成され、文書化されています。	必要に応じて、研究資料を安定した安全な施設に定期的に保管する計画が決定されています。	必要な研究資料が取得および/または作成および保管されている。
デジタル化 Digitization	必要に応じて、データ（および資料）をデジタル化する手順が作成され、文書化されています。	データ（および資料）のデジタル化に関する仕様のリストが作成され、文書化されています。	仕様リストのすべての項目は、他の研究者が理解できる言語で書かれています。	デジタル化されたデータ（および資料）を安定した安全なサーバーに定期的にアップロードする計画が策定されています。
データストレージ Data Storage	データ（およびデジタル化された資料）を保存するために必要な容量を見積もりました。	必要なストレージ容量を確保するための手配を行ないました。	データ（およびデジタル化された資料）が予想以上に増加した場合でも、十分なストレージ容量が利用可能です。	研究組織の構成員は、データ（およびデジタル化された資料）が十分な期間にわたって安全に保管されると確信しています。
物理的資料の保管 Storage of Physical Materials	必要に応じて、物理的（デジタル化された数値情報ではない）研究資料を保管するための設備、保管に必要な環境（温度、湿度など）、および保管にかかる費用を見積もりました。	物理的研究資料を保管するために必要な機器と施設を確保するための手配を行いました。	物理的研究資料の保管に必要な容量は見積もられており、計画以上に増加した場合でも対応可能です。	研究グループ構成員は、物理的研究資料が十分な長期間にわたって安全に保管されると確信しています。
ファイル形式 File Format	この分野で一般的に使用されている国際的に受け入れられたファイル形式を利用することが考慮されています。	データ解析を効率的に行うために、わかりやすいファイル形式を選択しました。	選択したデータ形式が文書化されています。	データは国際的に流通している形式でフォーマットされています。
バージョン管理 Version Control	複雑なデータを長期にわたって処理するためのバージョン管理計画を作りました。	データの最新バージョンを以前のバージョンから区別するための記述子を決めました。	以前のバージョンのデータを保存するためのアーカイブシステムを準備しました。	研究組織構成員は、既存のすべてのバージョンのデータが、それらにアクセスしたい他の研究者に利用可能になっていると確信しています。
バックアップ Backup	データをバックアップするためのさまざまなオプションが検討され、利用可能な資金あるいは組織のリソースが特定されました。	データを自動的にバックアップするためのリソースが利用できない場合は、手動バックアップの手順が用意されています。	起こり得る緊急事態に備えて、データを複数の場所またはリポジトリにバックアップするための手配が行われています。	研究者は、事故による損失の可能性に関係なく、研究データが利用可能であると確信しています。

前のスライドで赤枠で囲った部分の説明
(特にメタデータに関する説明)

メタデータとは

データに関する情報

(例) 何のデータか、データ取得場所・日時、記録されている内容(パラメータ)、データの精度やエラー、利用条件(ライセンス)、ファイル形式

必要性

データ処理・解析

データの公開(他人がデータの内容を理解し使うための情報)

メタデータおよびファイルフォーマットの一例

(IAGA-2002 format)

```
Format IAGA-2002
Source of Data Japan Meterological Agency
Station Name Memambetsu
IAGA CODE MMB
Geodetic Latitude 43.910
Geodetic Longitude 144.189
Elevation 42
Reported XYZG
Sensor Orientation HDZF
Digital Sampling 1.0 seconds
Data Interval Type Average 1-Minute (00:30-01:29)
Data Type Definitive
# D-conversion factor 10000
# K9-limit 350
# Format binary data.
# A complete set is available on the INTERMAGNET CD-ROM/DVD.
# Go to www.intermagnet.org for details on obtaining this product.
# CONDITIONS OF USE: The Conditions of Use for data provided
# through INTERMAGNET and acknowledgement templates can be found
# at www.intermagnet.org
# D conversion factor is a fixed value used to allow
# Declination to be converted from minutes of arc to equivalent
# nanoteslas. Set to H/3438*10000 where H is the annual mean
# value of horizontal intensity.
```

Metadata

DATE	TIME	DOY	MMBX	MMBY	MMBZ	MMBG
2020-06-16	00:00:00.000	168	25853.70	-4148.70	42592.90	.10
2020-06-16	00:01:00.000	168	25853.60	-4148.60	42592.60	.10
2020-06-16	00:02:00.000	168	25853.80	-4148.80	42592.30	.10
2020-06-16	00:03:00.000	168	25853.90	-4148.90	42592.10	.10

Data

データの解析 ANALYZING DATA

データの解析 ANALYZING DATA	初期 Beginning	発展 Developing	向上 Enhancing	完了 Completed
解析用ソフトウェア Software for Analysis	データ解析のために、必要な機能を備えた適切なソフトウェアを特定し、入手しました。	利用可能な既存のソフトウェアに研究に必要な機能の一部が欠けている場合には、それを補うために「自家製」ソフトウェアを開発しました。	利用予定のソフトウェアの解説書を入手するとともに、「自家製」ソフトウェアの説明文書を作成しました。	研究グループの構成員は、入手あるいは開発したソフトウェアが研究データの適切な解析を提供すると確信しています。
データの処理と解析 Data Processing and Analysis	すべてのデータ（および資料）はデジタル化の計画に従ってデジタル化されており、安全なサーバーに保管されています。	有意義な結果が得られるように、データは定性的および定量的に解析されています。	解析されたデータは、専門分野で一般的に使用される形式でフォーマットされています。	データの処理・解析方法やフォーマットを説明した文書を作成しました。
中間データ Intermediate Data	研究過程で生成された中間データが保存および保管に値するかどうかについて検討しました。	保存に値する場合は、中間データのフォーマットと保存に関する計画を立てます。	保持に値する中間データが取り出され、保存されます。	中間データに関する説明資料を作成しました。
データクリーニング Data Cleaning	データは視覚的にスキャンされ、ノイズ（エラー、スパイク、階段状のジャンプ等）と考えられる部分を特定・検出し、分類・整理しました。	（ノイズを処理し訂正された解析用のデータファイル作成には）元のデータファイルの複製が使われ、元のデータファイルは保持されます。	複製されたファイル内のすべてのノイズは分類され、削除あるいは修正されました。ノイズの分類方法が文書化されています。	ノイズ訂正済みのデータファイルを再検査し、間違っていると誤って判断されたデータが特定された場合は正しい元のデータに置き換えました。

前のスライドで赤枠で囲った部分の説明スライド
(特にデータクリーニングに関する説明)

中間データ

データを解析するためには、通常下記データクリーニングの他に、フィルターをかけたり、単なる電気信号から物理的なパラメーターに変換したりと様々な手を加える必要があり、その分、データとしての価値が高められている。また、他の人が再利用する場合にもオリジナルのデータよりも使いやすい。

- ・ 次の研究での再利用
- ・ 研究結果の確認
- ・ 論文査読段階でレフェリーから要求される可能性

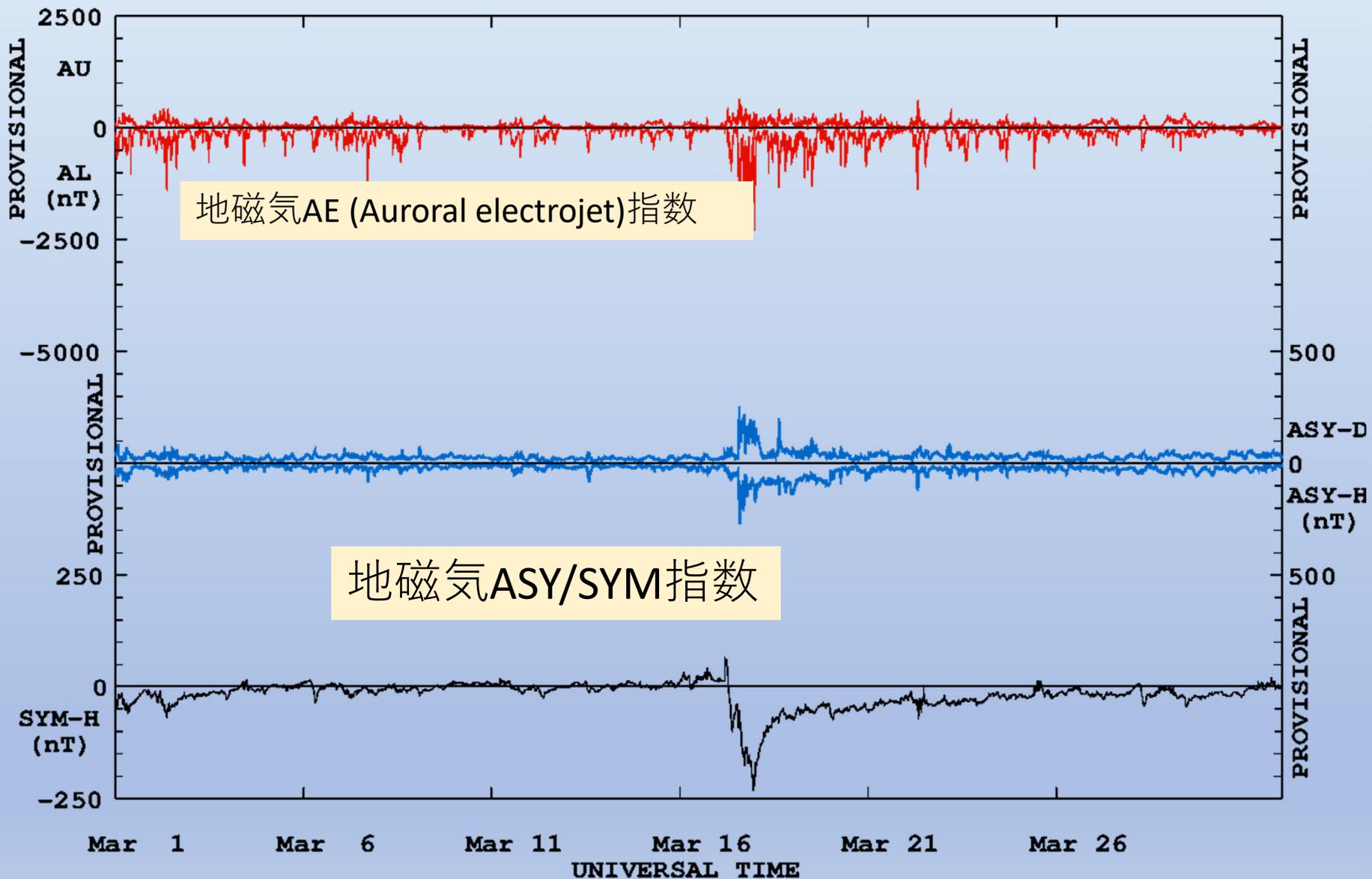
データクリーニング (ノイズ処理)

データ中のシグナル(有意な情報)とノイズ(不要な変動)の割合(S/N 比)によっては、解析結果が信頼できなくなるので、ノイズを可能な限り除去する必要がある。

INDICES

START: Mar 1, 2015 0000UT

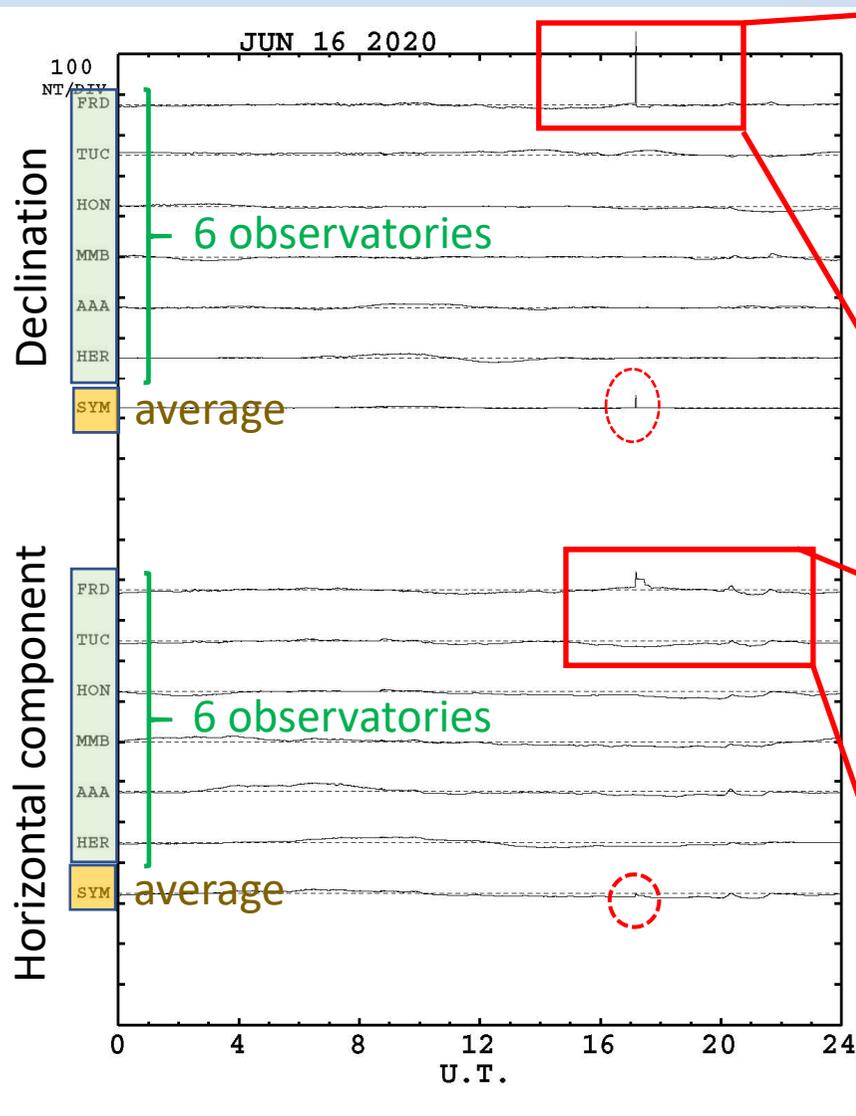
WDC FOR GEOMAGNETISM, KYOTO



データクリーニング(ノイズ除去)

Examples of noise in observed data

Geomagnetic data from 6 observatories



FRD: Fredericksburg, VA

noise

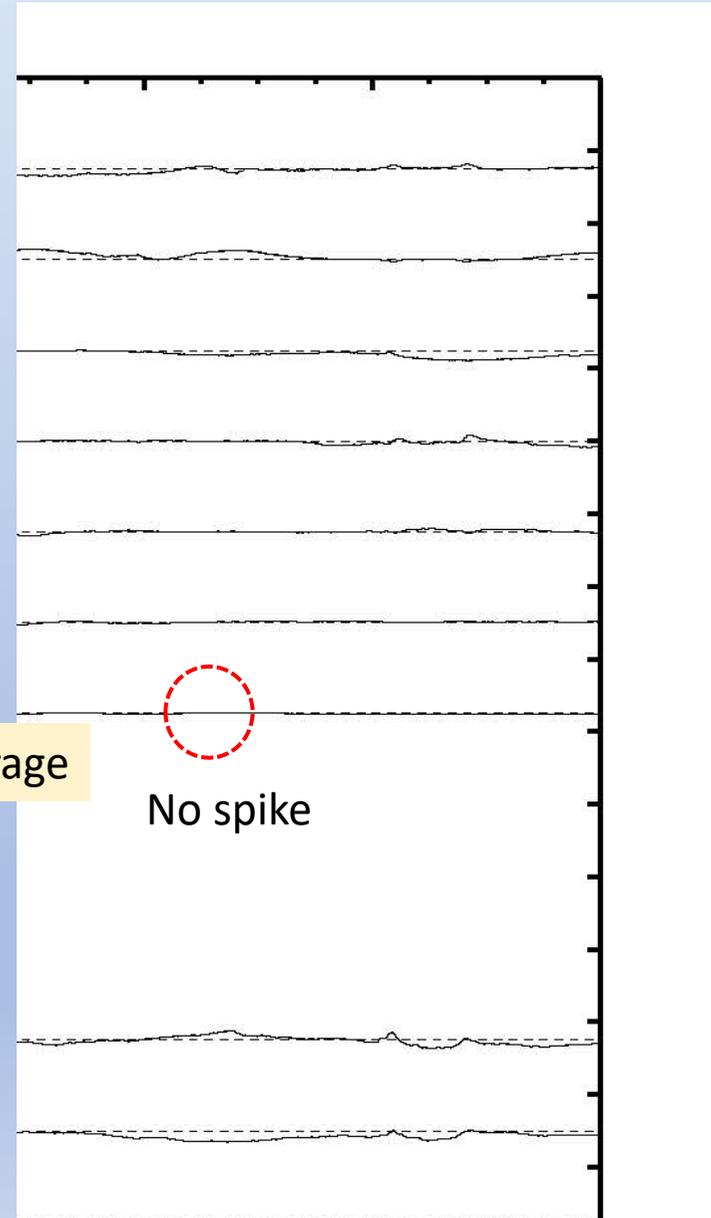
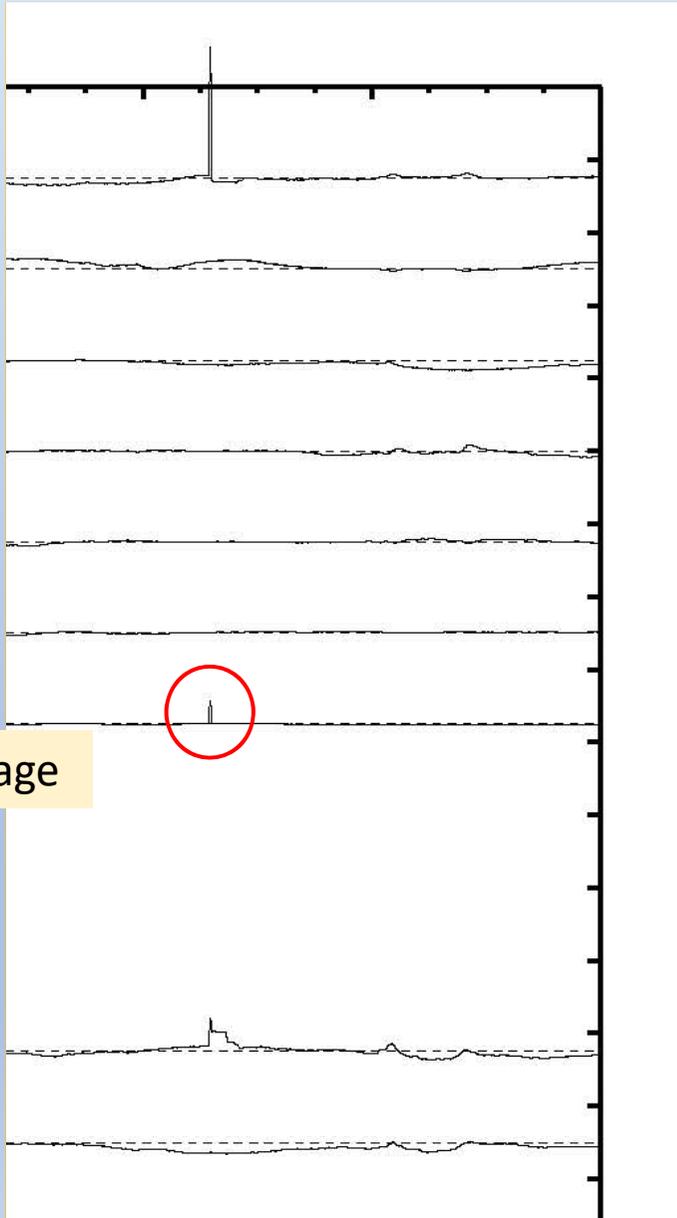
noise

real phenomenon

FRD

TUC: Tucson, AZ

Before A result of data cleaning After



データの共有または公開 SHARING OR PUBLISHING DATA

データの共有または公開 SHARING OR PUBLISHING DATA	初期 Beginning	発展 Developing	向上 Enhancing	完了 Completed
ライセンス Licensing	他の研究者によるデータの利用に関するルールを定めました。	国際的に認められた適切なライセンスを選択しました。。	選択したライセンスの使用要件は、研究グループ全構成員によって理解され、同意されています。	選択したライセンスの表示が公開データに添付されました。
メタデータスキーマ Metadata Schema	研究に必要なメタデータはどのようなものかを認識しており、その作成を容易にするために利用可能なソフトウェアを特定しました。	この分野で国際的に使用されているメタデータスキーマを選択しました。。	必要なメタデータとその重要性、使用方法、パラメータは、選択したメタデータスキーマに従って記述しました。	メタデータとそのスキーマが完全であり、他の研究者がデータ（およびマテリアル）を利用するのを支援する準備ができていると確信しています。
データセットへのマテリアルの収集 Collection of Materials into Datasets	研究資料をデータセットに編成するためのメタデータモデルが選択され、適応されています。	研究資料を説明するためのメタデータに必要な項目、データ型、ルールが定義され、文書化されています。	研究資料のメタデータ記述のルールは、データが保管されている機関のルールに可能な限り準拠します。正確な説明を保証するために URI (Uniform Resource Identifier) が使用される場合があります。	研究資料のメタデータ記述が、その分野で受け入れられているメタデータ形式と互換性がない場合、データセットの発見、アクセス、使用を効率化するために変換されます。
ドキュメンテーション Documentation	データ（および資料）の公開と共有に必要な文書の完全なリストが特定されました。	ドキュメントには、メタデータの構造を定義するスキーマが含まれています。	ドキュメントには、重要性、値、ファイル名、使用規則、物理量、データを取得するために使用した機器が含まれます。	必要なドキュメントはすべて作成され、データ（およびマテリアル）の公開と共有に伴うドキュメントファイルに追加されています。
データの完全性、有効性、信頼性、合法性、倫理 Data Completeness, Validity, Reliability, Legality, and Ethics	必要に応じて、研究におけるすべての実験および/またはヒト被験者の使用について IRB (治験審査委員会) の承認が得られており、研究に対する倫理的アプローチの必要性が満たされています。	この分野で認められているデータ収集および分析方法を使用することで、データ収集および分析の完全性（または誠実さ）、有効性、信頼性、合法性および倫理性が保証されています。	データ収集と分析の完全性、有効性、信頼性、合法性、倫理性の証拠が文書化されています。	データ収集と分析の完全性、有効性、信頼性、合法性、倫理性を証明するために、第三者によるレビューが実施されました。
研究資料の保存、返却、廃棄 Preserving, Returning, or Disposing of Research Materials	研究資料の保存、返却、または処分を選択肢は、そのために利用可能な資金も含めて検討しました。	研究資料の保存、返却、あるいは処分について決定し、これらの目的のための資金提供またはその他の取り決めが得られました。	研究グループ全構成員と研究者の所属機関の代表者は、研究資料の保存、返却、あるいは廃棄に関する決定と手順を理解し、受け入れています。	研究資料の保存、返却、処分が完了しました。
データの公開と共有 Publishing and Sharing Data	データ（およびデジタル化された資料）を公開および共有するための RDM 計画の検討を行いました。Web 上のオブジェクトにアクセスするためのデジタルオブジェクト識別子 (DOI) が取得されました。	データの所有権、ネットワークセキュリティ、バックアップシステムに関して、利用する可能性のあるデータリポジトリのポリシー、手順、コストを検討し、公開スケジュールを決定しました。	データ（およびデジタル化された資料）を公開および共有するために、適切なリポジトリを選択しました。資金は確保され、リポジトリとの調整も完了しました。	メタデータおよび文書を含むデータ（およびデジタル化された資料）が公開され、アクセスに関する以前の決定に従って他の研究者が利用できることを確認しました。。

前のスライドで赤枠で囲った部分の説明
(特にライセンスやメタデータスキーマに関する説明)

ライセンス

データ取得者およびそのグループに所属する研究者以外がデータを使う場合の条件を明確にするために、適切なライセンスを付与する。最近では、Creative Commons が提供している Creative Commons (C.C.) license がよく使われている。

メタデータスキーマ

メタデータを記述するための語彙や形式を定義したもの。その分野で広く使われているスキーマがあれば、それを使うことでデータ交換などを容易にすることができる。

クリエイティブ・コモンズ・ライセンス (CC Licenses)

<https://creativecommons.jp/licenses/>

<https://creativecommons.org/licenses/>

CC Licensesでは、データの利用条件として、

- (1) データ作成者へのクレジットの記述の必要
- (2) 非営利目的のみで使用可
- (3) データの改変禁止
- (4) 成果物を元のデータセットと同じライセンスを付与して公開する必要

の4つの項目の組み合わせで表示する。

C.C. ライセンス表示例 (<https://creativecommons.jp/licenses/>より)

CCライセンスの種類

作品を利用（再配布やリミックス作品の公開、実演等）するための条件は4種類あります。



表示



非営利

作品のクレジットを表示すること

営利目的での利用をしないこと



改変禁止



継承

元の作品を改変しないこと

元の作品と同じ組み合わせのCCライセンスで公開すること

これらの条件を組み合わせることができるCCライセンスは、**ぜんぶで6種類**。権利者は、自分の作品をどのように流通させたいかを考え、必要に応じて適切な組み合わせのライセンスを選ぶことになります。

組み合わせ例



これらのマークが表示されていることが、著作物にCCライセンスが付けられていることを示す目印です

(例1) Dublin Core metadata elements

要素タイプ名	定義とコメント
Title (タイトル)	情報資源に与えられた名前
Creator (著者・作者)	情報資源の創造に主たる責任を持つ人や組織
Subject (主題)	情報資源の主題あるいは内容を説明するキーワード
Description (内容記述)	情報資源の内容に関する説明や抄録
Publisher (公開者)	情報資源を現在の形で利用可能にした組織、出版社など
Contributor (寄与者)	「著者あるいは作者」以外で情報資源の創造に寄与した人または組織
Date (日付)	情報資源が作成された、または有効になった日付
Type (資源タイプ)	テキスト、イメージなど、情報資源の種類
Format (フォーマット)	情報資源のデータ形式
Identifier (資源識別子)	URI、ISBNなどの当該情報資源を一意に識別するための文字列または番号
Source (情報源)	情報資源を作り出す元になった別の情報資源に関する情報、出処
Language (言語)	情報資源を記述するために用いられている言語
Relation (関係)	他の情報資源との関係
Coverage (時間的・空間的)	情報資源の地理的または時間的特性
Rights (権利関係)	権利や利用条件に関する記述へのリンク

(<https://www.nrm.jp/column/column005>から引用)

(例2) 宇宙科学分野、日本ではIUGONETで使われている

“The Space Physics Archive Search and Extract (**SPASE**) Base Model”

The Resource Types currently in the [base data model](#) are:

<u>Annotation</u>	Explanatory or descriptive information that is associated with another resource.
<u>Catalog</u>	Listing of events or observational notes.
<u>Document</u>	Plain or formatted text that may include graphics, sound, other multimedia data, or hypermedia references.
<u>DisplayData</u>	A graphical representation of data.
<u>NumericalData</u>	Data product stored as numerical values in a specified format.
<u>Granule</u>	An accessible portion of another resource.
<u>Instrument</u>	A device that is used to sense and parameterize a physical phenomenon.
<u>Observatory</u>	The host (spacecraft, network, facility) for instruments making observations.
<u>Person</u>	An individual human being.
<u>Registry</u>	A location or facility where resources are catalogued.
<u>Repository</u>	A location or facility where resources are stored.
<u>Service</u>	A location or facility that can perform a well-defined task.

Each of these elements is broken down into sub-elements that provide sufficient detail for adequate description of data sets.

(<https://spase-group.org/data/>)

ドキュメンテーション

データの共有・公開だけでなく、論文作成における使用データについての正確な記述に有用。

データリポジトリの場合は、外部からの認証(例えばCore Trust Seal)取得に重要。

研究の各段階とルーブリックの使用目的・使用者

- **研究の計画段階**

研究計画の作成と予算獲得、主に研究代表者、ルーブリック全体を眺めつつ計画作成

- **研究の開始段階**

各研究者/チーム、使用データに関する確認、主にPlanning for Data, Organizing Data

- **研究の中間段階**

各研究者/チーム、研究の効率的推進、主にOrganizing Data, Analyzing Data

- **研究の終了直前**

各研究者/チーム、研究進捗状況の確認、論文作成
主にAnalyzing Data, Sharing or Publishing Data

- **研究の終了後**

研究代表者/チーム、取得データの整理、主にSharing or Publishing Data

まとめ

- ・ 研究推進に役立てるという観点から、研究の典型的な進捗段階に対応させて、ルーブリックを4枚の表に分けた。
- ・ 研究データマネジメントに関連して、知識として知っておくべき主な項目をルーブリックとしてまとめた。
- ・ 各自の研究の状況あるいは立場に応じて、ルーブリックに記載されている中から、必要な項目について自己評価する。
- ・ 研究計画の中のデータマネジメントプラン作成や、研究開始段階で特に考慮すべき項目等、使用目的に応じて使い分けることにより有効に利用されることを期待している。

予備スライド

	Ad Hoc	One-Time	Active and Informative	Optimized for Re-Use
Planning your project	When it comes to my data, I have a "way of doing things" but no standard or documented plans.	I create some formal plans about how I will manage my data at the start of a project, but I generally don't refer back to them.	I develop detailed plans about how I will manage my data that I actively revisit and revise over the course of a project.	I have created plans for managing my data that are designed to streamline its future use by myself or others.
Organizing your data	I don't follow a consistent approach for keeping my data organized, so it often takes time to find things.	I have an approach for organizing my data, but I only put it into action after my project is complete.	I have an approach for organizing my data that I implement prospectively, but it not necessarily standardized.	I organize my data so that others can navigate, understand, and use it without me being present.
Saving and backing up your data	I decide what data is important while I am working on it and typically save it in a single location.	I know what data needs to be saved and I back it up after I'm done working on it to reduce the risk of loss.	I have a system for regularly saving important data while I am working on it. I have multiple backups.	I save my data in a manner and location designed maximize opportunities for re-use by myself and others.
Getting your data ready for analysis	I don't have a standardized or well documented process for preparing my data for analysis.	I have thought about how I will need to prepare my data, but I handle each case in a different manner.	My process for preparing data is standardized and well documented.	I prepare my data in such a way as to facilitate use by both myself and others in the future.
Analyzing your data and handling the outputs	I often have to redo my analyses or examine their products to determine what procedures or parameters were applied.	After I finish my analysis, I document the specific parameters, procedures, and protocols applied.	I regularly document the specifics of both my analysis workflow and decision making process while I am analyzing my data.	I have ensured that the specifics of my analysis workflow and decision making process can be understood and put into action by others.
Sharing and publishing your data	I share the results of my research, but generally I do not share the underlying data.	I share my data only when I'm required to do so or in response to direct requests from other researchers.	I regularly share the data that underlies my results and conclusions in a form that enables use by others.	Because of my excellent data management practices, I am able to efficiently share my data whenever I need to with whomever I need to.

CDL (California Digital Library) がプロジェクトで出版点として蓄積にした
 担当プロジェクト

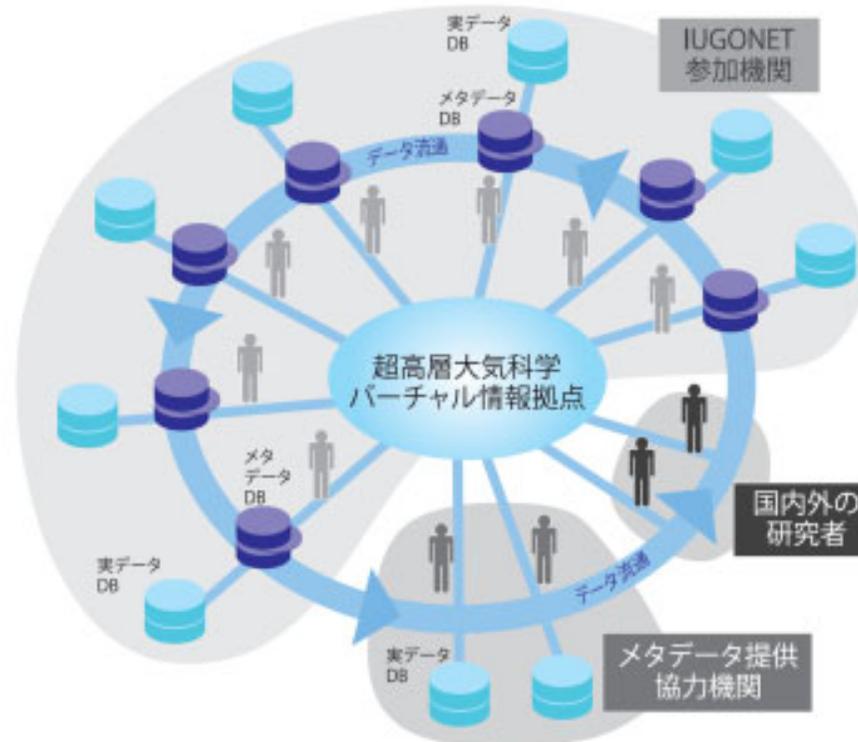
IUGONET: Inter-university Upper-atmosphere Global Observation NETWORK

IUGONETのとりくみ

IUGONETは、超高層大気関連の地上観測を長年推進してきた実績をもつ国立極地研究所、東北大学、名古屋大学、京都大学、九州大学の5機関の7組織が連携し、2009年に6カ年計画でスタートした大学間連携プロジェクトです。IUGONETは、次の2つのプロダクトの開発を柱としてプロジェクトを推進しています。

- ▶ IUGONETメタデータ・データベース (DB) → 5-6ページへ
- ▶ 解析ソフトウェアUDAS (iUgonet Data Analysis Software) → 7-8ページへ

これらの開発プロダクトは、2011年度末に正式版の公開・配布をプロジェクトのウェブページにおいて開始し、自由にご利用頂くことが可能です。今後も、参加機関内外からの地上観測データのメタデータの抽出とメタデータDBへの登録を促進するとともに、メタデータDB及びUDASの機能強化を実施します。



(IUGONETパンフレットから引用)



**CoreTrustSeal Trustworthy Digital Repositories
Requirements 2023-2025**

V01.00

<https://www.coretrustseal.org/>

物質科学用ルブリック - 研究データに関する計画 -

研究データに関する計画	初歩	基本	発展	完成
実験計画	実験の目的とその過程が明確にできる。 過程とは、材料生成のための原料の選定、利用する装置、評価する方法等	実験計画が研究ノートとしてまとめられる	研究ノートの書式が定められる	研究ノートの形式を随時改善できる
基本的データの種別と品質	研究を進める上で必要な基本的データを記述できる。 (例えば、原料のデータシート(CAS番号等)、装置のパラメータ、運転記録、解析データ)	取得すべき基本的データの種別や研究におけるそれらの役割が特定されている	取得できる基本的データの品質が特定されている。(例えば、原料の品質、試料の寿命、測定機器の精度、解像度)	基本的データの品質の変更、改善が管理できる(例えば、測定機器のバージョンアップや委託先の変更)
データ取得の方法	研究に必要なとなるデータの取得方法を決定した。	データ取得に必要なとなる装置や通信手段および記録装置を特定した。	データ取得装置や通信手段および記録装置導入の準備を始めた。	データ取得装置や通信手段および記録装置導入の準備が整った。
研究組織	研究データを取得し、研究を推進するメンバーが決まった。	研究を実施する各メンバーの持つ技能に応じて、役割が割り振られた。	メンバー同士が効率的に情報交換と相談を行い協力できる環境が用意され協力を開始した。	各メンバーは自身および他のメンバーの役割分担を理解し、共同研究をする用意が完全に整った。
研究資金	研究実施に必要なとなる研究費が見積もられ、申請を行う研究費配分機関を決めた。	各研究メンバーの人件費、装置や器具の購入、契約費用、出張などの申請額を決めた。	一つあるいは複数の配分機関への申請書類が整い申請した。	研究プロジェクトを遂行するためのデータ取得に必要なとなる十分な資金が確保された。
データの再利用	取得されるデータを再利用する場合は、データマネジメントの方法に変更が必要となるかもしれないということを理解している。	どのデータは再利用することができるか特定した。	データを再利用するために研究グループ外に公開するまでメンバーが優先的に利用する期間を決めた。	データマネジメント方法の変更も含め、データ再利用のための計画が完成し、研究グループ内で了承された。
データの公開・非公開	研究成果を発表予定の学術誌が要求しているデータ公開の条件を理解している。データ公開リポジトリでデータを共有する場合の選択要件も考慮した。	どのデータは公開し、どのデータは公開しないかを決めた。	データを公開する場合にデータセットにつけるライセンスと公開に使うリポジトリを決めた。	データの公開・非公開に関する様々な条件を文書化し、研究グループ全メンバーによって了解された。

人文・社会科学(地域研究)分野用ルーブリック - 研究データに関する計画 -

研究データに関する計画	初歩	基本	発展	完成
取得予定の資料	研究に必要となる基本的データの元となる資料を選定する。	取得すべき基本的データの種類および資料究の役割と関係を特定する。さらにデータの構築法について明確化する(目録作成か、デジタル翻刻か、資料の撮影は必要か、どのような分析を行うか等)。	基本的データの精度や品質および最低限カバーする期間や領域を特定する。	基本的データの精度や品質および最低限カバーする期間や領域を特定され、確実に取得する自信がある。
資料取得の方法	研究に必要となるデータの元となる資料の入手先と入手法を決定する。	資料の購入法・閲覧法・複製の可否、日本および資料収集地域における諸権利(著作権等)を特定する。	史資料入手のための準備を始める。	データ取得のための装置類・ソフトウェアの準備、および資料入手のための準備が整う。
研究組織	研究データを取得し、研究を推進するメンバーを決める。	研究を実施する各メンバーの持つ技能に応じて、役割を割り振る。	メンバー同士が効率的に情報交換と相談を行い協力できる環境が用意され協力を開始する。	各メンバーは自身および他のメンバーの役割分担を理解し、共同研究をする用意が完全に整う。
研究資金	研究実施に必要な研究費が見積もられ、申請を行う研究費配分機関を決める。	各研究メンバーの人件費、装置や器具の購入、契約費用、出張などの申請額を決める。	一つあるいは複数の配分機関への申請書類が整い申請した。	研究プロジェクトを遂行するためのデータ取得に必要な十分な資金が確保する。
データの再利用	取得されるデータを再利用する場合は、データマネジメントの方法に変更が必要となるかもしれないということを理解する。	どのデータを再利用できるかを特定する。	データを再利用するために研究グループ外に公開するまでメンバーが優先的に利用する期間を決めた。	データマネジメント方法の変更も含め、データ再利用のための計画が完成し、研究グループ内で了承する
データの公開・非公開	研究成果を発表予定の学術誌等が要求しているデータ公開の条件を理解している。データ公開リポジトリ等でデータを共有することを考慮する。	どのデータは公開でき、どのデータは公開できないかを決める。その際、日本および資料収集地域における諸権利(著作権等)の関係を考慮する	データを公開する場合にデータセットにつけるライセンスと公開に使うリポジトリを決める。	データの公開・非公開に関する様々な条件を文書化し、研究グループ全メンバーによって了解する。
永続的な資料・データ公開法(廃棄という選択肢も考慮する必要がある)	資料やデータ公開に関する研究機関、資金提供組織の要件を明確化する	研究終了後において、資料・データ公開に必要な設備・人員・経費等を見積もり、研究組織で行うか、外部委託するか等を決定する。	資料やデータ公開の永続的な公開法を文書化する。必要に応じて外部委託機関との交渉・調整を行う。研究組織内には研究代表者が属する研究機関において経費の永続的な負担法について交渉・調整を行う。	永続的な公開に向けた、外部委託機関および研究機関等との調整が終了し合意書を取り交わす。

学際研究用ルーブリック - 学際研究のチーム形成に関する計画 -

学際研究のチームに関する計画	初歩	基本	発展	完成
出会い	異分野交流会などに参加する	別分野の方と出会い、お互いの研究に対する思いを語る	学問を語る	連絡先を交換する
メンバーの選定	出会った人の中から一緒に研究をしたいと思ったメンバーをリストにする	どうしてその人と研究したいのか、何を深め合えると思ったのか、深め合いたいのか、理由を考え、言葉にすることができる		メンバーになってくれるよう打診する
連絡先	メンバーに誘いたい人と知り合いの人を探す	つなげてくれるようお願いする	話す機会を設ける	メンバーになってくれるよう打診する
刺激の共有	<ul style="list-style-type: none"> ・他者との対話 ・作品—美術館、映画、作品展 など ・他分野の研究会参加を週に1回以上行う 	同じ体験をメンバーと共有する	他分野で受けた刺激をどう持ち帰り、どうやってそれをシェアするかを考える	それぞれ参加して感じたもの、見えたものをシェアする
共有の方法	共有の方法に関して話し合う場を設ける	共有の方法に関して様々な方法があることをメンバーで認識をし、把握をしている。 <ul style="list-style-type: none"> ・話し合う・議論する・表現する・文字にする・公開研究会を行い、第三者に聞いてもらうことを意識するなど 	それぞれが得意で、安心することができ、伝わる共有の方法に関して話し合いを行う	それぞれの共有方法に関して定め、いつ共有を行うのかも定める
意識の共有	自分の今までの考えを話すことができる	相手の意見を聞くことができる	相手の意見の背景を理解することができ、相手の意見を受け入れることができる	相手の意見を
目的の共有	今回行う研究の目的をそれぞれ、考えてくる。	今回行う研究の目的を話し合った。	今回行う研究の目的をそれぞれの言葉で話すことができる。	今回行う研究の目的を文章にすることができ、全員が確認をした。
最終的な目標・成果公表	成果公表、最終的な目標に関して話し合う場を定期的に設ける	最終的な目標・成果公表にどのようなものがあり、可能性があるか、メンバーで把握をしている。 <ul style="list-style-type: none"> ・論文化・映像化・ワークショップ・啓蒙活動・授業・イベント など 	誰を対象とし、何をどう伝えたいのか、どう心に残って欲しいのか、何のために何をするのかをメンバーで話し合う	最終的な目標・成果公表方法を定めた
学際研究の認識	Inter, Multi, Transなどの違いを認識し、理解している。	InterやMultiではなく、Transになるためにはどのようなことが必要か話し合いを行なっている。	話し合いの結果、InterやMultiになってしまう要因を理解し、それを回避するために行う具体的な行動を理解している	Transな学際研究を実施する