

宇宙科学研究所のデータポリシー

2018年3月14日

1. 目的

本データポリシーは、宇宙科学研究所が研究開発を実施する際の、データの取り扱いに関する基本方針を述べたものです。

2. 適用

本ポリシーは、宇宙科学研究所が自ら取得・整備するデータに適用します。その他、データ作成に必要な装置、装置の搭載機会、資金、人的リソース等の一部を宇宙科学研究所が提供することによって他機関が取得・整備するデータについても、当該機関が、必要に応じて宇宙科学研究所と協力して、本ポリシーの趣旨を尊重して対応されることを期待します。

3. 本ポリシーが対象とするデータの定義

本ポリシーにおいて、「データ」とは、広い意味で科学的な価値を持つ情報であり、特定の物理的な媒体に依存せずに、汎用的・長期的に利用できるものを指します。汎用的な利用を想定していない個人的なメモや写真などの情報、研究グループの非公式のレポートや会議録、長期的利用を想定していない一時的な情報、物理的な実体としてのサンプルなどは、このポリシーの対象となるデータには含まれません。

以下に、ここで扱うデータの代表的な例と、簡単な説明を述べます：

- 源泉データ：衛星テレメトリなど、定まったフォーマットで判読されることが必要で、データ処理の源泉となるもの。これから適切なデータ処理を行うことによって、観測データや工学データなどを生成することができる。
- 観測データ：衛星、探査機、大気球、観測ロケットなどによって、天体や宇宙現象など、制御できない対象の物理的状態を測定した数値データ。多くの宇宙現象は顕著な時間変動を示すので、同じデータを再現できないことがある。
- 工学データ：衛星、探査機、装置の軌道、姿勢、温度など、データを取得する側の物理的状態を記述した数値データ。
- 実験データ：観測者が、対象に何らかの意図的な操作を加えることによって測定した数値データ。多くの場合、実験を繰り返すことにより、同じデータを再現することができる。
- シミュレーションデータ：観測データ・工学データ・実験データなどを模擬するために、計算機によって生成されたデータ。計算を繰り返すことにより、同じデータを再現することができる。
- 探査機で得られた小惑星サンプルや微小重力実験で得られたサンプルの分析結果を定量化した数値データや、探査機が測定した天体の形状を再現した数値データなど。分析・再現精度が向上することによって、改訂されることがある。

- 汎用的・長期的な利用を意図して作成された、デジタル化された書類、写真、画像、映像など。

また、各データを記述するためのデータ(メタデータ)や、データを利用するために必要なツール、ソフトウェア、アルゴリズム、説明文書等も、対応するデータに準じて扱います。

4. 非公開データ、公開データの考え方

宇宙科学研究所は、本ポリシーが対象とするデータについて、個別の状況に応じて、データ毎に「非公開」とするか「公開」とするかを定めます。宇宙科学研究所は、論文や学会等で発表された結果を再現するために必要なデータ等、科学的成果のエビデンスとなるデータを公開します。それ以外のデータについても公開を原則としますが、以下の場合に限って、データを「非公開」とします：

1. データを公開すると、個人情報保護や公共安全等に支障がある場合。
2. データ処理が不完全であることが明示されておらず、データを公開すると間違った結果が発表される可能性がある場合（注1）。
3. データの取得や作成を行った研究チームが、一定期間占有利用する場合。
4. データの利用権を特定他機関等に付与する場合。

宇宙科学研究所は、原則として、非公開データの存在を、それを非公開とする理由とともに、明示します。ただし、データの存在を公開しないことにメリットがあると考えられる場合には、その存在を公開しない場合もあります。非公開データについては、利用範囲と非公開期限を定め、その後、公開に移行するか、非公開を続けるか、あるいは廃棄するかを判断します。

衛星、探査機、大気球、観測ロケットなどによる観測データについて、装置チームが機器較正を行ったり観測提案者が占有利用したりするための非公開期間が必要な場合、その目安はデータ取得後から約1年です。

（注1）装置のパフォーマンス等を例示する目的で、データ処理が不完全でも、精度が保証されないことを明記したうえでデータ公開する場合もあります。データ処理の問題で公開できない場合には、問題を解決して公開に至るまでのスケジュールを明示します。

5. 公開データのポリシー

宇宙科学研究所は、公開データが広く利用されることが科学の進歩につながるとの信念に基づき、以下の方策を実施します：

1. 公知の知識のみで公開データを利用できるように、適切なデータ処理やデータの説明を行います。
2. 公開データは、利用できる状態で、長期間（注2）保管します。
3. 必要な公開データを簡単に見つけ、使いやすくするためのサービスを無償で提供します。
4. 永続的な識別子（注3）を用いるなどして、公開データを引用しやすくします。

(注2) データが利用価値を有する期間に対応し、最低30年とします。

(注3) 代表的なものに、DOI (Digital Object Identifier) があります。

6. 公開データ利用の際のルール

公開データを利用するには、以下のルールに従ってください。なお、このルールは政府標準利用規約(第2.0版)(注4)に準じたもので、クリエイティブ・コモンズ表示 4.0 国際 (CC BY 4.0) と互換性があります：

1. 公開データは、原則として、営利目的、非営利目的を問わず、複製、送信、加工も含め無償で利用できます。ただし、一部のデータについては、宇宙科学研究所以外の第三者が用途を制限している場合があります
2. データの利用に関しては、「宇宙航空研究開発機構・宇宙科学研究所」または「ISAS/JAXA」と出典を明示してください。さらに、データによっては、その取得・整備・公開等に関わった個人や組織が出典の明示を要求している場合もあるので、利用者の責任において、それに従ってください。
3. 公開データを加工して利用する場合には、加工したという事実を明記するとともに、できる限り、どのような加工を行ったかを示してください。
4. 宇宙科学研究所は、利用者がデータを用いて行う一切の行為について何ら責任を負うものではありません。

(注4) http://www.kantei.go.jp/jp/singi/it2/densi/kettei/gl2_betten_1.pdf

7. データ保存の考え方

宇宙科学研究所は、原則としてすべてのデータを可能な限り保存しますが、コストやリソースの観点から、データを廃棄せざるを得ない場合もあります。以下に、データ保存についての考え方を示します。

- 公開データは、長期間保管する。
- 原理的に再現できないデータは、長期間保管する。
- データ処理によって観測データや工学データなどを再現するための源泉データは、できるだけ再処理が可能な状態で、長期間保管する。
- 原理的に再現できるが、再現に大きなコストがかかるデータは、できるだけ廃棄しない。
- データ処理、実験、計算等によって比較的容易に再現できるデータは、廃棄する可能性がある。

以上