

望ましい科学データ公開・利用のあり方について

宇宙航空研究開発機構(JAXA)

宇宙科学研究所(ISAS)

科学衛星運用・データ利用センター(C-SODA)運営委員会

2013年12月9日

はじめに

当委員会では、先に「科学データの公開・利用について」と題するレポートを策定し¹、その中で、JAXA の衛星・探査機が取得した科学データの公開・利用に関する原則と指針を示した。本文書は、その続編として、JAXA の衛星・探査機による科学データの公開・利用に関する現状を分野横断的に調査し、共通の問題点を抽出した上で、それらの問題点を解決し、JAXA における望ましい科学データの公開・利用を実現するための方向性を提案するものである。

本文書で用いる用語は「科学データの公開・利用について」の定義に従っているが、本文書だけを読んで理解できるように、頻繁に用いる用語を以下に説明する。

- データセンター

科学研究機関において、長期アーカイブを担当する部署。宇宙科学研究所では C-SODA がそれに対応しているが、本文書は、望ましい科学データ公開・利用のあり方一般について述べたものであるため、本文中では C-SODA という固有名称の使用は避けている。

- アーカイブ

最終的な公開を前提としたデータ処理(プロセッシング)、データ保存、データサービスの集合体。

- データプロダクト

なんらかの科学観測の結果得られるデータを、利用しやすいように処理(プロセス)した生成物。

- データ処理(プロセッシング)

データプロダクトを生成するためのアルゴリズムを確立し、それを適用する作業。

- プロジェクト

一般的に、プロジェクトとは企画、研究計画、事業開発計画のことであるが、JAXA において「プロジェクト」とは、ミッションを達成する手段として設定され、特定の資源と時間のもとで時限的組織により実施する有期的活動のことを言う(付録 2 参照)。両者を区別するために、本文書では後者

¹ <http://c-soda.isas.jaxa.jp/bibliography/KagakuData.pdf>

を「プロジェクト」と表記する。

- 短期アーカイブ

進行中の衛星「プロジェクト」が、主に「プロジェクト」の遂行に必要な情報とデータプロダクトを管理・利用するために開発するアーカイブ。

- 長期アーカイブ

衛星「プロジェクト」終了後に、「プロジェクト」関係者以外でもデータプロダクトの長期利用を可能とするためのアーカイブ。

目次

第一部：科学データ公開・利用に関する現状分析と分野共通の課題	5
1.1. 現状分析	5
1.2. 分野共通の課題	5
第二部：課題解決のための提案	9
2.1. 衛星・アーカイブの開発と「プロジェクト」審査	9
2.2. アーカイブ審査と審査の視点	10
2.3. アーカイブ審査の体系について	13
付録1：分野ごとの背景と課題	15
1.1. 高エネルギー天文学・赤外線天文学	15
1.2. 月惑星科学	16
1.3. 太陽物理学	16
1.4. 太陽地球系物理学(STP)	16
1.5. 宇宙環境利用科学	18
付録2：ミッション、「プロジェクト」、フェーズ移行審査について	21
付録3：宇宙科学研究所における科学衛星審査について	22
付録4：「プロジェクト」終了後のアーカイブ開発・運用体制について.	24
執筆者一覧:	26

第一部：科学データ公開・利用に関する現状分析と分野共通の課題

1.1. 現状分析

当委員会では、本レポートの策定に先だって、JAXA の各衛星・探査機が取得した科学データの公開・利用状況(アーカイブ化)の現状分析を行った。衛星、観測機器、データ種別毎に以下の 7 つの項目と判定基準を設定し、これらの基準が満たされているどうかを判定した。

1. 「ポリシー」：ミッション当初に設定したデータポリシーにおいて、データプロダクトの作成（プロセッシング）から、保存、公開、長期アーカイブまでが考慮されているか？
2. 「全データ処理」：全取得データが適切にプロセッシングされているか？
3. 「内部文書」：担当者が替わっても、それを元にしてアーカイブ開発・運用を引き継ぐことを可能とする内部文書が存在するか？
4. 「長期保存」：科学的価値のあるデータプロダクトが、利用可能な状態で長期保存されているか？
5. 「公開」：作成されたデータプロダクトが公開されているか？
6. 「サービス」：ユーザーがデータプロダクトを利用するために必要なサービスが整備されているか？
7. 「ユーザー向けデータ記述文書」：特定のソフトウェアに依存せずにそのデータを扱えるレベルのデータ記述文書が存在するか？

これらの現状分析から分野ごとの課題を抽出し²、さらに、そこから分野共通の課題を抽出した³。本レポートでは個々の衛星や観測機器についての判定結果は提示しないが、以下に、抽出された分野共通の課題を示す。それらの課題の解決のための提案を、第二部に示す。

1.2. 分野共通の課題

A) アーカイブ対象となるミッションの偏り

² 付録 1 参照。そこでは、各分野のデータ公開・利用の現状を理解する際の参考となる、分野ごとの背景も述べた。

³ 解析ソフトウェア、データ解析マニュアルの整備等、ユーザーサポートの現状については、本レポートのスコープを超えると判断し、分析をおこなわなかった。

ISAS で古くから衛星を打ち上げてきた科学分野においては、必ずしも十分でないにしても ISAS のリソースを用いてアーカイブが整備されているが、それに比べて新しい分野では、整備が遅れていることが多い。テレメトリデータについても、ISAS の衛星については長期保存されている（SIRIUS データベース）が、それ以外の「プロジェクト」や「プロジェクト」化されていないもの（たとえば ISS の観測装置）については長期保存が保証されていない。JAXA の衛星・探査機が取得した科学データの中でも、アーカイブ対象となるものには偏りがあり、それを改善することが分野共通の課題である。

B) データポリシーの設定

どの分野においても、当初からデータ公開や利用のポリシーが定められていた「プロジェクト」（あるいは観測機器やデータプロダクト）については、アーカイブ整備が比較的進んでいる。一方、全データ処理、長期保存、公開がうまくいっていない、あるいは高次データプロダクトにユーザーの意見が反映されていないと思われるケースについては、ほとんどの場合、その背景にデータポリシーの欠如がある。多くの場合、打ち上げ前のデータポリシーには長期アーカイブの整備までは含まれていないが、長期的な視点に立ったデータ公開・利用のポリシーを、「プロジェクト」初期の段階から定めておくことが重要である。その際、データは誰のものかという観点よりも、「科学データの公開・利用について」で述べたように、「科学データは人類共通の知的資産として広く公開して活用されるべきという原則」に基づいて、データポリシーが定められるべきである⁴。

C) データの利用可能性の判断

データフォーマットや機器に関する情報が失われているために実質的には利用不可能なデータや、データ処理の源泉ではないので、再処理の際もアクセスする必要がないテレメトリデータが長期アーカイブ化され、その整備に継続的な運用コストが投じられている場合がある。保存するデータについては、使えるデータと使えないデータを仕分けておくこと、また、使用不可能なデータに

⁴一方、地球観測分野においては高次データに関しては高次データを作りだした者がその権利を持っている。公開・配布についても自由に決められる。また、商業利用をした場合はロイヤリティを取るなどの縛りがある。

については、いつまで保管しておくべきかを明らかにしておくことが課題である。

また、電波干渉計データなど、地上で取得したデータと相関を取って初めて意味を持つ衛星データもある。そのような、単独では利用できないデータのアーカイブ化の方針を定めておくことも重要である。

D) JAXA と他機関、および JAXA 内の組織間の関係の整理

JAXA 外の大学やデータセンターでデータ処理が行われたり、データが発信されたりしている場合、その継続性や長期保存が保証されていないことが多い。たとえば、装置の PI が所属する機関がデータ処理やデータ公開を行っていて、予算がつかなくなった後のデータの扱いが決まっていないケースがある。また、JAXA 外に構築されたデータセンターのミラーとして JAXA からデータを再発信したり、データセンターの運用が終了した後に JAXA がデータを引きとる必要が生じたりする場合もある。そのような場合に責任が曖昧にならないよう、機関間でインタフェースと役割分担を明確にしておくことが必要である。

また、一つの衛星について、多機関でデータ処理やデータ公開が行われていて、全体像が見えにくくなっている場合がある。全体像をユーザーに明らかにした上で、複数機関を活用した積極的なデータ公開・利用とサービスの多様化が実現されることが望ましい。

さらに、JAXA 内においても、衛星の計画や運用、データ処理・公開に関わる組織は複数存在するので、組織間の責任分界を明確にしておくことが重要である。

E) 文書整備と短期アーカイブから長期アーカイブへの移行

担当者が替わってもアーカイブ開発・運用を引き継げるようにするための開発者向け文書や、特定のソフトウェアに依存せずにそのデータを扱えるレベルのユーザー向け文書が整備されていないことが多い。短期アーカイブ担当の「プロジェクト」関係者が「プロジェクト」を離れた後、情報が文書として残されていないと、長期アーカイブへの移行は困難である。そのために、「プロジェクト」関係者向けの短期アーカイブは完成しても、「プロジェクト」外のユーザーが利用可能な長期アーカイブが完成しない場合がある。また、長期アーカイブが完成した場合でも、ユーザー向け文書が整備されていないと、長期間に亘ってアーカイブから科学的成果を産出し続けることは困難である。

F) リソースの不足とシステム化の遅れ

以上の課題が認識され、アーカイブ整備のために何を実施すべきかが明確であっても、人員、予算等のリソース不足のために、整備が進まないことがある。各研究コミュニティにおいて、既存衛星の運用と並行して次期衛星の装置開発を行っている場合が多く、あらたな衛星開発に限られたマンパワーが割かれると、既存衛星のデータ処理・公開がおろそかになりがちである。特に、衛星プロジェクトの後期になると、重点的に次期衛星にリソースが割かれ、長期アーカイブの完成まで至らない場合が多い。新たな衛星開発や運用とは独立に、すでに取得したデータのアーカイブ化に専念できる専門知識を持った人材とそのため予算確保が分野横断的な課題である。

また、アーカイブの構築が特定の個人あるいはPIチームの能力に依存していることが多い。アーカイブ開発の制度やシステムの整備によって、できるだけ個人あるいはPIチームの能力に依存せずにアーカイブ開発を進められるようにすることが必要である。その際、限られたリソースによるアーカイブ開発の効率化を図るため、できるだけプロジェクトや分野を超えて利用可能なソフトウェアやツールの基盤整備を進めることが望ましい。

第二部：課題解決のための提案

ここでは、前節で抽出された分野共通の課題解決に向けた提案を行う。

2.1. 衛星・アーカイブの開発と「プロジェクト」審査

「科学データの公開・利用について」で述べたように、そのままでは科学目的に利用できないテレメトリデータを、使いやすい形にまで高次処理し、さらに長期間にわたって保存し、広い範囲の研究者によって使われてこそ、科学的成果が最大化される。すなわち、衛星のデータ処理・保存・サービスが一体となった「データアーカイブ」は衛星のサブコンポーネントと同様の重要性を持つものとして位置づけられるべきである。

衛星「プロジェクト」の実行段階は、付録 2 の通り、時間的に区切られた多くの連続する「フェーズ」からなる⁵。次のフェーズに進む、あるいはフェーズ内の次の段階に進むためには、有識者によるフェーズ移行審査が実施され、それを通過することが必要である。その後、「プロジェクト」終了審査を経て、「プロジェクト」終了となる。

アーカイブ開発や運用においても、「プロジェクト」の各フェーズに対応したフェーズが存在すると考えられ、その状況は審査対象とするべきである。たとえば、「ミッション定義審査」の段階で、新天体を発見したらそのデータは即時公開する等のデータポリシーを定義しておくことが必要である⁶。それに応じて、「システム要求審査」では、データポリシーを満たすためのシステム要求（新天体のデータを観測後 1 時間以内に世界に配信する、等）を示すことが必要である。「システム定義審査」では、データプロダクトを作成・配布するためのシステム（たとえば、観測後 1 時間以内にデータを処理してウェブで配信するシステム等）が定義されているかどうか、審査対象となる。基本設計審査、詳細設計審査を経た後に、アーカイブシステムの製作・試験を行う。同様に、打ち上げ後の「プロジェクト運用延長審査」や「終了審査」においても、アーカイブ開発や運用の状況が審査対象となるだろう。

このように、「プロジェクト」のフェーズ移行に必要とされている各審査において、衛星や装置の開発・運用状況の審査だけでなく、アーカイブ開発・運用

⁵ ISAS の科学衛星に関しては、付録 3 も参照のこと。

⁶ NASA/ESA では、通常、観測機器の提案書の段階で、データポリシーやデータ処理の議論が含まれている。

状況の審査が含まれることが望ましい。

2.2. アーカイブ審査と審査の視点

前節で述べたように、「プロジェクト」の各審査にデータアーカイブの視点を入れることは重要である。その目的は、「科学データの公開・利用について」で述べた原則(以下に再掲)が満たされ、かつ 1.2 節で述べたデータ公開・利用に関わる分野共通の課題が解決されていることを確認することである。

「科学データの公開・利用について」で示した三原則は以下の通りである：

1. **データプロセッシングの原則**：すべての科学データについて、機器較正やデータ処理アルゴリズムを適用し、公知の知識だけでそこから科学的成果を引き出せるような段階に至るまでの処理（プロセッシング）を行う。
2. **データ保存の原則**：取得したすべての科学データは、使用できる状態で永久に保存する。
3. **データサービスの原則**：データセンターは、データプロバイダを明らかにした上で、そのデータが長期にわたってできるだけ広い範囲のユーザーに使われるようにするための基盤サービスを無償で提供する。

また、1.2 節で述べた分野共通の課題 A-F に対応し、アーカイブ審査において確認すべき項目は以下の通りと考えられる：

- A) 科学データを生み出す衛星「プロジェクト」はすべて審査の対象となっていること。すなわち、JAXA 内の組織に依らず、科学データを生み出す衛星、探査機、あるいは観測装置すべてについてアーカイブ審査が行われること。
- B) 三原則に沿ったデータポリシーが適切に設定されていること。
- C) データプロダクトの利用可能性が明確化されていること。
- D) JAXA と他機関、および JAXA 内の組織間の関係が整理されていること。
- E) 短期アーカイブから長期アーカイブへの移行を可能とする内部向け文書、や長期アーカイブ運用のためのユーザー向け文書が整備されていること。
- F) 適切なリソース（人員、予算）が確保され、アーカイブ開発が個人に依存せず、システム化されていること。

以下で、アーカイブ開発・運用に関して、各段階で必要と考えられる審査と、そこで要求される「審査の視点」(確認すべき項目)を説明する。なお、衛星本体や

検出器の開発とアーカイブ開発のフェーズは必ずしも一致しておらず、以下で説明するアーカイブ審査と、「プロジェクト」の各段階で実施されるフェーズ移行審査(付録 2,3)との対応は自明ではない。よって、どの段階の「プロジェクト」審査でどのアーカイブ審査を行うかは、状況に応じて決められてしかるべきである。

1. データ利用ポリシーに関わる審査

三原則に従ったデータ利用ポリシーが策定されていることを確認する(確認項目 B)。具体的には、「プロジェクト」において、どれがアーカイブ化すべきデータか、誰がどのデータを使う権利があるのか、データの優先期間はどうか⁷、を明確化する。また、公開しないデータに関しては、その理由(公開しない理由、メリット)を示す⁸。当審査は、衛星「プロジェクト」の初期の段階に実施されることが望ましい。

2. 短期アーカイブ設計に関わる審査

主に「プロジェクト」関係者向けの短期アーカイブの設計を審査する。テレメトリからプロセッシングを経て、具体的にどのようなデータプロダクトを生成するのか、各データプロダクトの意味、想定寿命、利用可能性、公開範囲等を確認する(確認項目 C)。全体的なデータポリシーに加え、データプロダクト個別のポリシーが必要であれば、それを定める(確認項目 B)。データ処理、公開、利用に関わる機関(組織)間の役割分担と責任分界点を明らかにする(確認項目 D)。さらに、短期アーカイブ開発および運用に必要なリソース(人員、予算)が確保されていることを確認する(確認項目 F)。また、短期アーカイブ設計が文書化されていることを確認する(確認項目 E)。衛星の「プロジェクト」化が決定した時点およびそれ以降に複数回実施され、徐々に設計が詳細化されていくことが望ましい。

⁷多くの科学者にとって、自分で開発した衛星や観測装置を用いて新たな発見を行うことが衛星「プロジェクト」実行の強いモチベーションであるので、「プロジェクト」チームに優先的にデータ使用权が与えられることは当然認められる。しかし、優先期間の後にはデータは原則公開されるべきである(ただし、次の脚註参照)。

⁸工学データ等に関しては、技術的な優位性を保持するために、公開が難しい場合もある。そのような非公開データに関しては、機密性を保持することによって、守るべき価値が存在することを明らかにするべきである。

3. 短期アーカイブ開発完了に関わる審査

設計に基づき、短期アーカイブの開発が完了したことを確認する。衛星の打ち上げ前に実施されることが必要である。

4. 短期アーカイブ運用進捗に関わる審査

定常観測運用開始後に実施される。「プロジェクト」に関わる分野の研究者がデータ利用を可能にする文書が整備されていること（確認項目 E）、適切なリソースが投入され（確認項目 F）、データプロセッシング、データ保存、データサービスが、三原則に従って進行していることを確認する。仮に、リソースの不足によってデータ公開・利用が進んでいない場合には、その増強を提言する可能性もある。また、新たなデータプロダクトの作成等により新たなデータポリシー設定が必要になった場合には、それを確認する（確認項目 B）。

科学衛星に関しては、定常運用期間終了後に、運用延長審査を実施する機会が多い(付録 3)。それに即して当審査を行うことが望ましい。また、仮にすでに取得したデータの公開・利用の状況が芳しくない場合は、運用延長に警告を出す場合もありうる。

5. 短期アーカイブから長期アーカイブ移行のための審査

主に「プロジェクト」が開発する短期アーカイブから、主にデータセンターが担当する長期アーカイブへの移行がスムーズに行われることを確認するために実施する。短期アーカイブからデータプロダクトを取捨選択し、長期アーカイブとして何を残すか(確認項目 C)、「プロジェクト」とデータセンターの役割分担ははっきりしているか、「プロジェクト」終了後にデータセンターが恒久的に長期アーカイブの運用を行えるようになっているか（確認項目 D）、データセンターが長期アーカイブ運用や「プロジェクト」外のユーザーサポートを可能とするような文書が整っているか（確認項目 E）、長期アーカイブ運用のための人員・予算等のリソースが確保されているか(確認項目 F)、を確認する。これによって、プロジェクト終了後に、第三者が長期アーカイブを使って科学的成果を産出できることを確実化する。

科学衛星に関しては、「プロジェクト」終了審査に即して当審査を行うことが望ましい。ただし、「プロジェクト」終了までにデータ処理が完了せず、短期ア

アーカイブが完成していない場合もある。その場合には、「プロジェクト」が終了しても、短期アーカイブ完成まで高次データプロダクト作成を継続するための組織を新たに定義するべきである（付録 4 参照）。その場合でも、短期アーカイブ完成後に、長期アーカイブ移行のための審査は必要である。

2.3. アーカイブ審査の体系について

上記で、データアーカイブの開発や運用についても、衛星の開発や運用と同様、各フェーズにおける審査が必要であることを述べた。ここで、アーカイブ審査は、その目的が異なる複数の審査からなることに注意する必要がある。以下で、アーカイブ審査の体型を明確化する（図 1）。

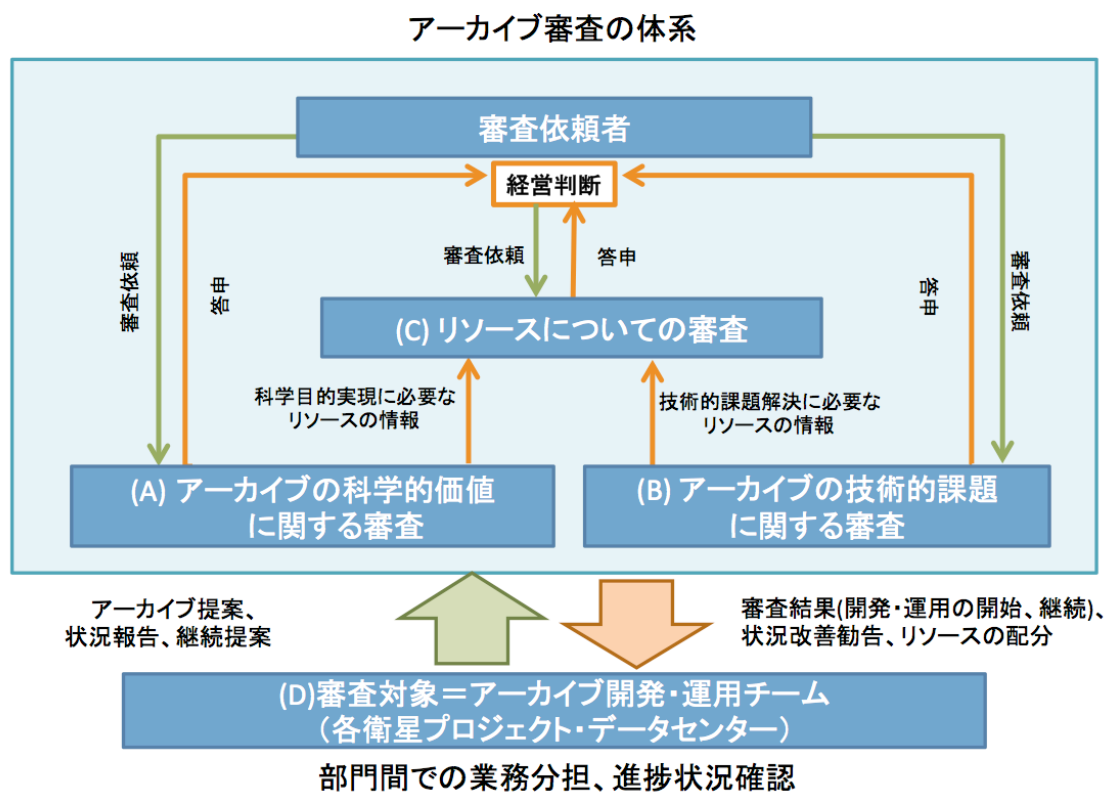


図 1：アーカイブ審査の体型

図 1 のアーカイブ審査体系を構成する、各部分について説明する：

A：アーカイブの科学的価値に関する審査

- データのアーカイブ化の必要性和科学的価値を確認する。
- 科学目的を達成するために、どのプロダクトをアーカイブ化すべきか、優先順位も含めて判断する。また、そのために必要な、科学的専門性を

伴うリソース(ポストク等)を見積もる。

- その分野における科学的見識の深い審査員が実施する。

B：アーカイブの技術的課題に関する審査

- 上記 A の審査のあとに実施する。A で科学的価値を有すると判断されたプロダクトのアーカイブ化に関して、技術的な実現可能性を評価する。
- アーカイブ化の実現に、どれだけのタイムスケールが必要か、また、技術的な問題を解決するために必要なリソース(エンジニア等)を見積もる。
- その衛星の分野におけるデータの取り扱いや、衛星の技術的側面を良く理解している審査員が実施する。

C：リソースについての審査

- 上記 A、B の審査の後に実施する。これらの審査の結果を踏まえ、経営的な視点からリソースの配分について判断する。
- 研究機関が実施する他の事業の状況も理解し、それを考慮した上で経営的な判断ができる審査員が実施する。

D：審査対象=アーカイブ開発・運用チーム

- アーカイブ開発・運用計画を、上記 A、B、C の審査に提出する。
- 短期アーカイブを担当する衛星「プロジェクト」と、長期アーカイブを担当するデータセンターが協力して計画を策定する。
- 審査に際し、データセンターが衛星「プロジェクト」からデータプロダクトを受け入れる、あるいはデータサービスを引き継ぐ際の条件等、役割分担を明確化しておく。

上記 A、B、C の審査は、必ずしも別々に実施する必要はないが、その目的と審査員に期待される知見や見識が異なることを認識しておくことが必要である。また、アーカイブ審査を構成する審査依頼者、審査対象、各審査の主体(審査委員会)が、具体的にどの組織や職位に対応するかは、状況によって異なる。しかし、実際の審査にあたっては、どの組織や職位がどの部分を担当するかを明確化しておく必要がある(付録 3 参照)。

付録 1: 分野ごとの背景と課題

ここでは、宇宙科学における各分野（高エネルギー天文学・赤外線天文学、月惑星科学、太陽物理学、太陽地球系物理学、宇宙環境利用科学）の科学データ公開・利用状況を理解する際に参考となる、分野ごとの国内外の背景と、現状分析から抽出した分野ごとの課題を述べる。

1.1. 高エネルギー天文学・赤外線天文学

1980年代になって、ほとんどの天文観測データがデジタル化されるに伴い、デジタルデータのアーカイブ化が進んできた。特に、異なった天文台の間でデータを容易に交換できるよう、FITS(Flexible Image Transportation System)フォーマットが考案されると、それが短期間に世界中の天文台のデファクトスタンダードになった。現在では、ほぼすべての地上観測データ、天文衛星データが FITS フォーマットで記述されている。

FITS フォーマットを扱う解析ソフトウェアパッケージの標準化が進んでいて、同じソフトウェアを使って、複数のミッションや装置のデータ解析を行うことができる。ソフトウェアパッケージについては、米国の研究機関で開発されたものが使われることが多く、そのほとんどが無償でサポートされている。

高エネルギー天文学に関しては、NASA/GSFC の HEASARC がアメリカを始めとする世界の高エネルギー天文衛星データの集約を進めていて、HEAsoft と呼ばれる汎用解析ソフトウェアのサポートを行っている。1993年に打ち上げられた「あすか」衛星のデータ処理、アーカイブ化は宇宙研と NASA/GSFC が協力して実施した。「あすか」が高エネルギー天文衛星では初めて全面的に FITS フォーマットを採用した衛星であり、それ以降の高エネルギー天文衛星もすべてそれとよく似た FITS フォーマットを使っている。あすか、すざくに続き、ASTRO-H のデータ処理、アーカイブ化も NASA/GSFC と協力して進められている。

赤外線天文学分野においても FITS フォーマットが採用されており、汎用的なソフトウェアを利用することが可能である。カリフォルニア工科大学に併設されている Jet Propulsion Laboratory (JPL)に所属する Infrared Processing and Analysis Center (IPAC)が、NASA の赤外線アーカイブデータ（地上観測

データを含む)を集約している。

1.2. 月惑星科学

月惑星科学の分野では、NASA が科学データアーカイブを行うために開発した Planetary Data System (PDS) というシステムが広く用いられている。また、探査機の位置、速度、姿勢、視野などが科学データとは別にアーカイブされており、その基盤となるのが、JPL が開発した SPICE システムである。

また、JAXA も参加している International Planetary Data Alliance (IPDA) がデータ共有のために開発した国際標準プロトコル PDAP (Planetary Data Access Protocol) が存在し、今後の月惑星科学データは PDAP で配信することが標準になっていくと考えられる。また、月惑星科学は地球観測との関係も深く、地球観測で標準的に用いられている GIS (Geographic Information System) の応用もなされている。

デジタルデータだけでなく、「はやぶさ」や「はやぶさ 2」のように、天体からサンプルを持ち帰って研究できることが、月惑星科学分野の際だった特徴である。サンプルの分析結果であるデジタルデータに関しては、原理的に、衛星や探査機の観測データと同様に扱うことが出来る。また、サンプル自身のアーカイブ化はデジタルデータのアーカイブとは別の重要課題である。

1.3. 太陽物理学

天文学と同じく、FITS フォーマットが広く用いられている。ほぼすべての衛星で、原則として観測データは即時公開である。データ解析は「ようこう」の解析システムを起源とする、IDL 上で動作する”SolarSoft”が、標準的に用いられている。SolarSoft のソースコードは公開されているが、リビジョン管理されておらず、過去のバージョンによる結果まで遡ることは困難である。ただし、常に最新版による結果が尊重され、再現性はあまり重要視されていない。

1.4. 太陽地球系物理学(STP)

太陽地球系科学の研究対象は、地球における中間圏から太陽圏まで広範囲に及ぶが、日本の科学衛星が主に対象としてきたのは、電離圏、磁気圏など、宇宙プラズマ現象に関わるものである。宇宙機によって現象の発生する「その場」において観測を行っていることに最大の特徴があり、時系列データの取り扱い

が中心となる。宇宙プラズマ現象のその場観測として、電磁場計測、プラズマ粒子観測、撮像観測が標準的であるが、中性粒子やダストカウンターや電波遮蔽観測も STP に含まれる場合がある。

研究者によるデータの使い方として、個別イベントの詳細解析と、長期間にわたるイベントの統計的解析がある。一点観測だけでは時間変動と空間変動を区別できないので、関連する多くの周辺情報（他の宇宙機や地上データ）をできるだけ多く集めることが求められる。

データアーカイブの利用者は、「プロジェクト」チームの内外を問わず、全世界の STP 研究者である。短期アーカイブ=L1 データが対象、長期アーカイブ=L2 データ以降が対象と考えられる⁹。データアーカイブに求められる機能として以下が上げられる：イベント発見の効率化、イベント母集団データ収集の効率化、データ解析作業の効率化、解析ソフトウェアの効率化、周辺情報収集の効率化等。

当分野の特徴として、小規模の観測装置が多く、観測機器チームがデータプロセッシングをおこなうが、十分な体制を組めずに長期間の運用に耐えられないということが挙げられる。このような場合、アーカイブを整備可能かどうかは PI 次第である。また、ミッションが長期にわたる場合があり、当初メンバーが現役を離れると、軽微な修正すら困難で大幅なデータ処理体制の遅れにつながる可能性がある。

また、較正作業の難しさから、膨大なデータのすべてについて完全に品質を保証できず、リクエストベース、部分的にしか十分較正されない場合がある。よって、公開データの品質については、それが保証できる低レベルに設定されてしまう場合がある。

データフォーマットに関しては、国内では、今までは各観測機器チームが SDB (Science Data Base) などの独自バイナリーフォーマットを採用してきたが、STP 分野の国際標準である CDF(Common Data Format)への移行が順次進められている。

⁹ データプロダクトレベル定義については、「科学データの公開・利用について」の付録を参照。太陽地球系物理学分野では、L1 データが PDS レベルで言う Edited Data に、L2 データが、Calibrated Data 以降(Resampled Data, Derived Data を含む)に対応する。

1.5. 宇宙環境利用科学

主に国際宇宙ステーション(ISS)の与圧部(きぼう実験棟)を用いて実施されている宇宙環境利用科学研究は、材料科学、燃焼科学、流体科学、宇宙生命科学から構成される。これらの実験結果の種類は2つに大別される。

一つ目は、実験試料(サンプル)自体を種々の分析、解析に供するものであり、多くの場合、最終的に実験試料は細切れ状態となる。分野としては、細胞培養、植物などの宇宙生命科学分野、半導体などの結晶成長、液体拡散などの材料科学分野に多い。なお、植物の生育実験、半導体結晶成長実験などでは、一定時間ごとに温度-時間のデータや植物生長のビデオ観察などがある。

二つ目は、流体科学、燃焼科学分野に多く存在する現象のビデオ画像記録である。宇宙実験における流体実験では、主として流れ場の観察のためトレーサ粒子を入れ、ビデオにより観察し、同時に温度あるいは温度差を計測している。さらに、流れ場が3次元になった場合には3軸方向からのビデオ観察を同時刻に合わせて流れ場を決定する必要がある、それには高度な技術を必要とする。

宇宙環境利用科学には多種多様なデータが存在し、そのうちアーカイブされるべきものは、主としてビデオ観察(温度データなどが整合したもの)である。ここでは、更にその詳細を見ることとしよう。

流体科学

H3~H26年まで実施されたあるいは、今後予定されている流体実験は、気泡の運動、マランゴニ対流実験と2相流実験である。前者2つは小型ロケットによって実施され、ビデオ画像が実験結果である。これらのビデオ画像は解析後、DVDにして保管している。1つの画像でおよそ1程度であり、5実験が行われている。

また、H18年以降JEMで実施されたマランゴニ実験が、現在までに3実験ある。1実験あたり数日の実験でこれらの総容量は15TBほどの録画量と成っている。しかし、実験目的として注目しているのは20分程度の部分である。また、2相流実験、マランゴニ実験がH22~26年にかけて行われる。これらの総録画量は、25TBほどと見積もられる。

データアーカイブ化するには、先にも述べたが温度、流れ場時刻の同一性の担保した3次元解析が必須である。

材料科学分野

H3~H32年まで実施あるいは、予定されている実験は、溶液成長実験3テーマと結晶成長実験及び、浮遊炉の熱物性計測及び、新機能材料創成である。溶液成長実験は、ビデオ画像がすべてであり、3実験を合わせると5TBほどなる。本実験では、成長の過程と温度の相関が必須であり、このための整合を取る必要がある。

静電炉による熱物性実験は、比較的短時間で、準備も入れて20分程度であろう。特に、粘性係数、表面張力などの熱物性計測は、ビデオ画像が生命線であり、質の高い画像が望まれる。本実験のH32年までの画像総容量は3TBほどと見込まれる。

なお、アーカイブ化は、比較的容易であり、励起振動あるいは、減衰中の試料形状の精度良い測定がすべてと成る。また、新機能材料は、地上に持ち帰り、各種の解析を行うことから、最終的には試料は、粉々の状態と成る。従って、アーカイブ化の必要はなく、発表論文の結果のトレースとならざるを得ない。

燃焼分野

燃焼分野における実験のほとんどが、ビデオ画像である。一部には、NO_x量の分析などが無いわけではないが、微小重力実験においては、ほとんどが、OHラディカルの発光、火炎の形状、進展形態などを含めてビデオ画像が、すべてと言える。

H3~H26年までの実施済み及び実施予定の燃焼テーマは、4テーマほどあり、このうち2テーマは小型ロケット実験により実施された。これらの実験結果は、すべてビデオ画像である。容量的には、両者あわせて1TBほどであるが、火炎伝播と温度変化の相関が必須である。

また、今後予定されている燃焼実験も、画像が実験結果である。JEMでは、体系的な実験実施が、可能であることから、実験回数も多くなり、両者あわせて3TBほどの容量が必要である。アーカイブ化に当たっては、液滴の位置情報と火炎伝播の相関、電線の火炎伝播挙動実験では、流れる電流と火炎伝播挙動の相関の整理が必要である。

宇宙生命科学分野

宇宙生物科学分野には、重力生物学、植物学、細胞生物学及び、放射線生物

学の 4 課題が存在する。

- 植物実験

植物学実験では、主に微少重力下での生育速度、根の屈曲などの画像取得が 1 つの実験結果である。この場合、成長速度は、流体などの流れ場の変化とは、比べものにならないほど小さいため、1 実験で 100MB ほどである。H3~H26 年まで実施済み及び、実施予定の植物実験は 4 テーマあり、このうち 2 テーマは実施済みである。これらの画像では温度、湿度、時間の整合が必須である。

- 重力生物学分野、細胞生物学

重力生物学としての実験は、線虫が、代表的な例である。H16~H26 年の間で実施済み及び、実施予定のテーマは、4 テーマであり、このうち 2 テーマが、既に実施されている。

これらの実験では、現在のところ画像取得はなく、地上に持ち帰った試料を分析して成果を創出している。また、細胞生物学の実験も画像取得はなく、地上での遺伝子解析などの分析が主体である。このため、アーカイブ化は、論文でのデータのトレースとなる。

- 放射線生物学

放射線生物学の実験も高エネルギー粒子が、当たってもすぐには変化が現れないことから、in-situ の観察は行われていない。地上に持ち帰った試料について遺伝子の異常性、新たな遺伝子の発現などを解析している。このため、アーカイブ化では、然したる作業はない。

付録 2: ミッション、「プロジェクト」、フェーズ移行審査について

JAXA プロジェクトマネジメント規定および実施要領によると、「ミッション」とは、JAXA が行うべき特定の計画・活動をいい、それは以下で定義する「プロジェクト」の終了時において、あるいは終了後の利用・研究等を経て、最終的に獲得すべき状態または成果を含む。

「プロジェクト」とは、ミッションを達成する手段として設定され、特定の資源と時間のもとで時限的組織により実施する有期的活動のことを言う。

大規模で複雑なシステムを確実にかつ効率的に開発するため、「プロジェクト」は、フェーズに区分して進めるものとし、「プロジェクト」マネージャは、各フェーズで行うべき作業内容を定義した上で、フェーズ移行審査における次フェーズへの移行可否判断を踏まえ、順次これを進めるものとする。が、JAXA が定義する「プロジェクト」フェーズとフェーズ移行審査の関係である。

また、フェーズの段階を示す以下の用語も用いられる：フェーズ A(概念設計、計画決定)、フェーズ B(基本設計)、フェーズ C(詳細設計)、フェーズ D(製作・試験、射場整備)、フェーズ E (初期運用、定常運用)、フェーズ F(後期利用)。

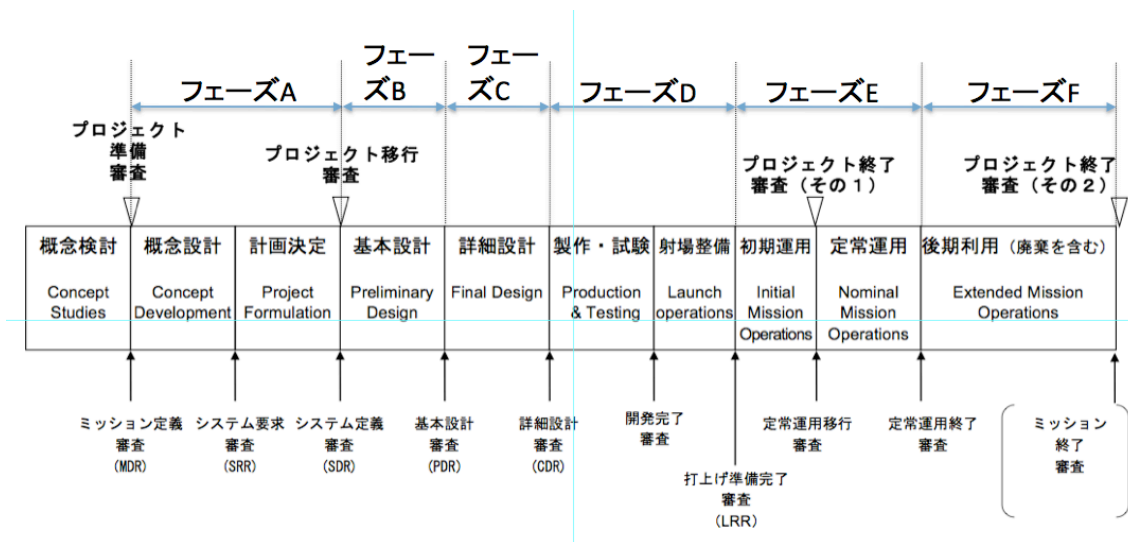


図 2: 「プロジェクト」のフェーズと審査の関係
(プロジェクトマネジメント実施要領から)

付録 3: 宇宙科学研究所における科学衛星審査について

宇宙科学研究所における科学衛星プロジェクトの実施は、JAXA 統合以前からの経緯等から、付録 2 で述べたフェーズの定義に完全に従うわけではなく、以下の特徴がある。科学衛星プロジェクトを立ち上げたい有志は、宇宙理学委員会下で数年間のワーキンググループ(WG)活動を行う。WG 活動で十分にミッション提案の準備ができたとき当事者が判断し、なおかつ宇宙科学研究所から「プロジェクト」の公募が出た場合に、WG は理学委員会にプロジェクト審査を申請する。本委員会の下に「プロジェクト審査委員会」が設置され審査に当たり、結論を本委員会に報告する。理学委員会で認められた場合にはさらに上部委員会である運営協議会で審議し、宇宙科学研究所の「プロジェクト」として選定する。

ここで正式に「プロジェクト」(フェーズ B)となり、その後、基本設計審査(PDR)を経てフェーズ C、詳細設計審査(CDR)を経てフェーズ D となり、衛星の製作に入るのは JAXA 方式と同じである。理学委員会による「プロジェクト」審査は、JAXA の MDR と SRR を同時に行い、さらに委員会として宇宙科学ミッションたり得るとの判断を確認するプロセスが含まれていると解釈できる。

科学衛星においては、定常運用期間を終了しても、「プロジェクト」が観測運用の継続を希望する場合がほとんどである。その際、理学委員会下に設置される「観測運用延長審査委員会」が審査を行い、結論を本委員会に報告する。それが承認された場合には、理学委員会は宇宙科学研究所に運用延長を勧告する。通常、運用延長期間は三年間であり、「プロジェクト」が更なる観測運用継続を希望する場合には、再度、観測運用延長審査が行われる。運用延長審査においては、取得したデータが適切に公開・利用されているかどうかも審査対象になる(2.2 節参照)。

「プロジェクト」終了審査に関しては、JAXA のモデルケースでは、「その 1:投入した資源や体制が適切であったか?」、「その 2:ミッションの目的、求められる役割を果たしたか、波及効果は得られたか?」からなる。科学衛星については、前者は宇宙科学研究所が、後者は理学委員会が実施する。理学委員会下の終了審査では、取得したデータのアーカイブ整備状況(科学的価値の高いデータプロダクトが作成され公開されているか?)も、審査対象になる(2.2 節参照)。

2011 年 11 月に観測運用を終了した「あかり」衛星については、2012 年度に、理学委員会および ISAS によって複数の終了審査が実施された。理学委員会下の終了審査(審査委員、委員長は理学委員)において、科学的価値の高いデータプロ

ダクト作成が完了していないため、あかり「プロジェクト」終了後も、高次データプロダクト作成のためのデータ処理を継続することが勧告された(図 1 の A に相当)。その後、ISAS の下で終了審査が二回開催され(担当部署は宇宙科学プログラムオフィス、審査委員長は宇宙科学プログラム・ディレクタ)、その中でデータアーカイブ計画の技術的妥当性の評価(図 1 の B に相当)とそれに要するリソースの正当性の評価(図 1 の C に相当)が行われ、了承された。それを受けて、あかり「プロジェクト」終了後、2013 年度から新たに「あかりデータ処理・解析チーム」が設置され、データ処理を継続することとなった。

付録 4 : 「プロジェクト」終了後のアーカイブ開発・運用体制について

一般に、衛星「プロジェクト」が短期アーカイブを開発し、「プロジェクト」が終了するに伴い、データセンターが長期アーカイブを引き継ぐ。しかし、「プロジェクト」が終了した時点で短期アーカイブが完成していない、すなわち、高次データプロダクトを作成するために、データプロセッシングを継続することが必要な場合がある。そのためには、衛星「プロジェクト」終了後に、データプロセッシングを専門的に行い、それを予算、人事等の面において管理するための新たな組織（「データプロセッシングチーム」）を、データセンターと独立に定義することが望ましい。このように短期アーカイブ業務と長期アーカイブ業務を独立の部署で実施することによって、組織としての責任分界が明確となり、短期アーカイブから長期アーカイブへの移行が確実化される。

アーカイブに必要な作業	データプロセッシング		データ保存	データサービス (配布、検索等)
プロジェクト実施中の担当	プロジェクト (必要に応じてデータセンターの計算資源を利用)		データセンター またはプロジェクト	データセンター またはプロジェクト
プロジェクト終了後の担当	↓ 完了している場合 なし	↓ 継続する場合 新たなデータプロセッシングチーム	↓ 短期アーカイブ ↓ 長期アーカイブ ・データセンター担当部分は継続 ・プロジェクト担当部分は取捨選択してデータセンターに移行	↓ 短期アーカイブ ↓ 長期アーカイブ ・データセンター担当部分は継続 ・プロジェクト担当部分は終了またはデータセンターに移行

図 3 「プロジェクト」実施中と終了後のアーカイブ作業担当

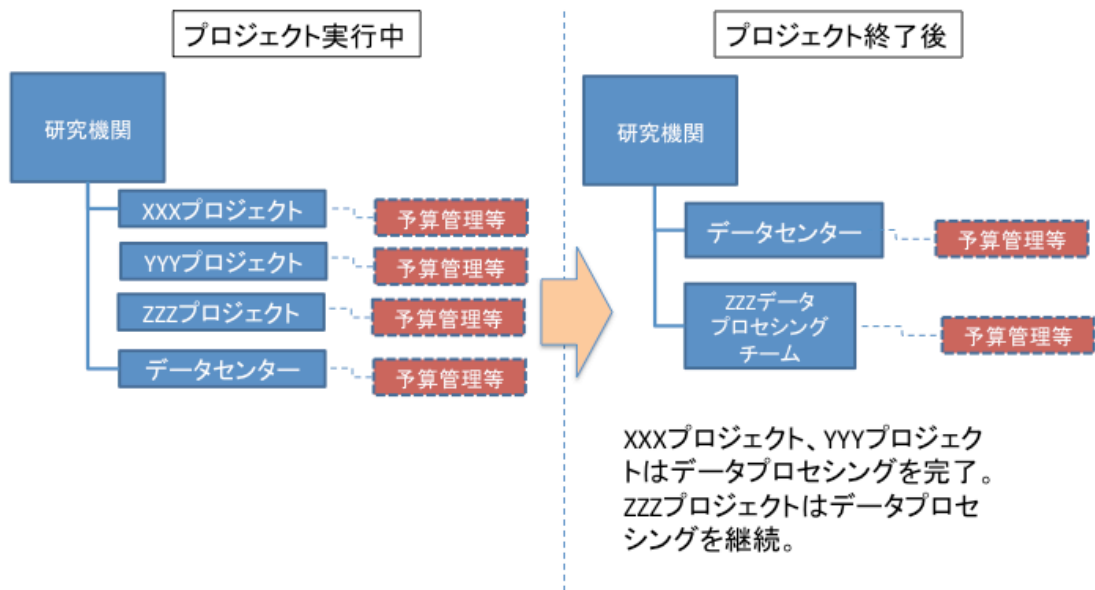


図 4 「プロジェクト」実行中と終了後の組織体制

執筆者一覧:

H21-24 年度科学衛星運用・データ利用センター運営委員(50 音順、敬称略)

石岡 憲昭(宇宙科学研究所; H24 年度)

稲谷 芳文 (宇宙科学研究所)

海老沢 研 (宇宙科学研究所)

尾中 敬 (東京大学)

鎌田 幸男 (宇宙科学研究所; H21-23 年度)

阪本 成一 (宇宙科学研究所)

篠原 育 (宇宙科学研究所)

清水 輝久(宇宙科学研究所;H23-24 年度)

菅原 正行(宇宙科学研究所; H24 年度)

常田 佐久(国立天文台; H24 年度)

出村 裕英 (会津大学)

中村 正人 (宇宙科学研究所; H21-23 年度)

波形 寿英 (宇宙科学研究所;H21-22 年度)

福田 徹 (JAXA 地球観測研究センター)

松崎 恵一 (宇宙科学研究所)

水本 好彦 (国立天文台)

村田 健史 (情報通信研究機構)

山内 茂雄 (奈良女子大学)

山本 幸生 (宇宙科学研究所)

横山 央明 (東京大学)

依田 眞一 (宇宙科学研究所; H21-23 年度)