

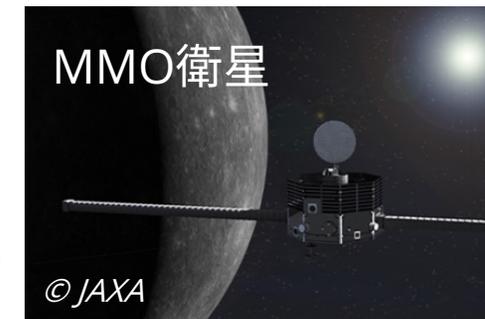
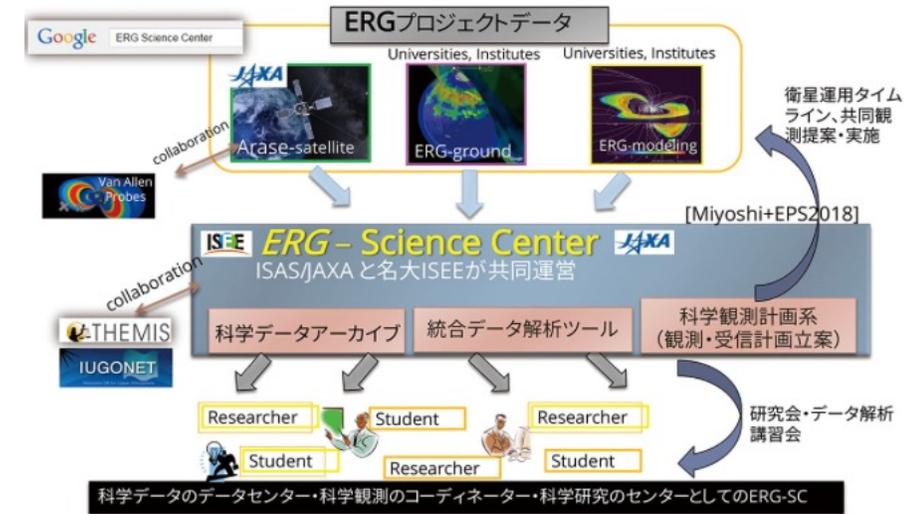
# IUGONETメタデータの来し方行く末

堀 智昭 (名古屋大ISEE)

# 堀のIUGONETとの関係

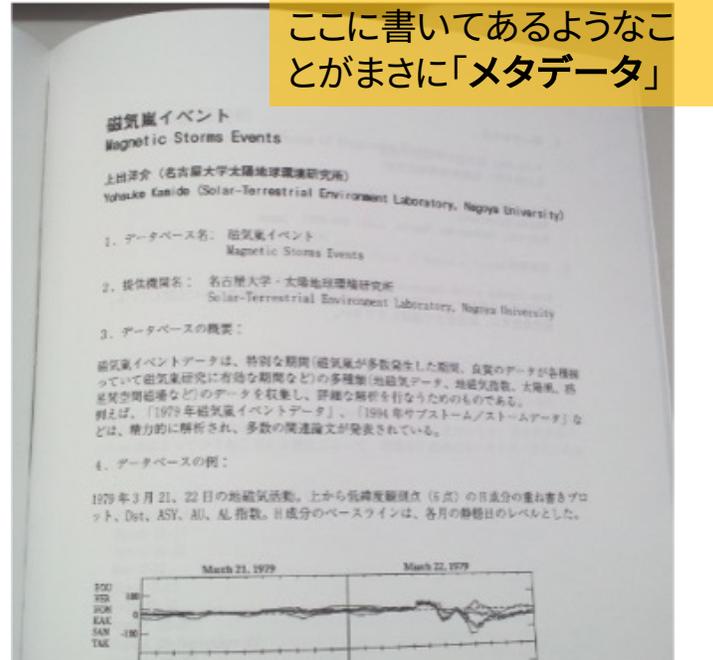
- ▶ 初期のIUGONET開発メンバー (2009-2013) メタデータ策定サブグループのリーダーだった。
  - ▶ SPASEグループとの交渉, IUGONETフォーマットの策定
  - ▶ XML schemaの準備、メタデータ作成手引きの作成
- ▶ その後、名大ISEEのERGサイエンスセンターで **あらせ衛星及び連携地上観測データのデータ (ファイル・アーカイブ・解析ツール) 開発。**
- ▶ 最近では太陽圏サイエンスセンター(CHS)で **日欧水星探査ミッションBepiColomboの衛星データ開発も。**

ERGプロジェクトとERG Science Center (ERG-SC)



# データベースカタログからメタデータDBへ

## 従来型のデータベースカタログ



## オンラインアクセス・検索が可能なメタデータ・データベース

IUGONET Web Service  
Upper Atmosphere x IDL x Web Technology

Type-A

UDAS web Available!

IUGONET Metadata DB

Home  
IUGONET MDB Search Help

Browse Data  
Entire Data / Resource  
Resource Type

Browse Service  
Browse Service

UDAS  
Iugonet Data Analysis Software

IUGONET

All

Free Word:  
Free Word  
(e.g. ion)

Time:  
from YYYY-MM-DD

Data Types:  
Data Set (  Nu

**IUGONET DataSet**

| Instrument Type   | Observed Region                          | ERG Campaign                                |
|---|--|---|
| Satellite:<br><input type="checkbox"/> AKEBONO              | <input type="checkbox"/> CHAMP           | <input type="checkbox"/> COSMIC             |
| Ground-Based:<br><input type="checkbox"/> SMART (Telescope) | <input type="checkbox"/> DST (Telescope) | <input type="checkbox"/> FMT (Telescope)    |
| <input type="checkbox"/> Geomagnetic Indices                | <input type="checkbox"/> Magnetometer    | <input type="checkbox"/> WDC Geomag., Kyoto |
| <input type="checkbox"/> Geomag., Kakioka                   | <input type="checkbox"/> MM210           | <input type="checkbox"/> AWS                |
| <input type="checkbox"/> EA Radar                           | <input type="checkbox"/> LT Radar        | <input type="checkbox"/> MU Radar           |
| <input type="checkbox"/> MW Radar                           | <input type="checkbox"/> WP Radar        | <input type="checkbox"/> X-Band Radar       |
| <input type="checkbox"/> GPS Receiver                       | <input type="checkbox"/> Na-Lidar        | <input type="checkbox"/> EISCAT             |
| <input type="checkbox"/> VHF Radar                          | <input type="checkbox"/> VLF/ELF         | <input type="checkbox"/> Ionosonde          |
| <input type="checkbox"/> Others                             |  |   |

Keyword:

Timespan:  To  [Set Detail](#)

Information  
ASCII Downloader was added: 19 Dec. 2019.

▶ 収集された情報は膨大であり価値があるが、それらを有効に活かすことができない

- ▶ 冊子の中を検索できない
- ▶ 印刷物ゆえに更新できない
- ▶ 計算機やWeb経由でアクセスできない
- ▶ 破損・紛失する可能性

- レコードを検索できる
- 内容を更新できる
- Web経由でアクセス可能
- 他のシステムにデータを引き継ぎ可

# メタデータスキーマの比較 (2010年当時)

IUGONET  
Metadata DB for Upper Atmosphere

## メタデータフォーマットの比較

| 項目                   | IPY MDP, DIF | STARS, ISTP                  | SPASE                                | PDSラベルフォーマット     |
|----------------------|--------------|------------------------------|--------------------------------------|------------------|
| やり取りする時のフォーマット       | XML          | XML                          | XML                                  | テキスト             |
| どの科学分野、どのDBが使っているか   | IPYDID, GCMD | STP分野<br>STARS-DB,<br>CDAWeb | STP, Heliospheric<br>VHO, VMO, ViRBO | Planetary<br>PDS |
| 数値データセットを説明する要素の有無   | ○            | ○                            | ○                                    | ○                |
| カタログデータセットを説明する要素の有無 | ○            | ○                            | ○                                    | △                |
| 非デジタルデータを説明する要素の有無   | △            | ×                            | ○                                    | ○                |
| データファイルを説明する要素の有無    | ×            | ○                            | ○                                    | ○                |
| 可視化属性の有無             | ×            | ○                            | ○                                    | △                |
| 必須(null不許可)の要素数      | 8+18         | 21+72+18+26 (DB全て)           | 15+7+7+6+3                           | 8(+20)           |
| 新しい要素、単語の追加が容易       | ×            | △                            | ○                                    | △                |
| ドキュメント、公開ツールが揃っているか  | △            | △                            | ○                                    | ○                |

2010/2/2 Hori, T., et al., IUGONET全体会議 4

- ▶ 当時超高層大気分野, space physics 分野およびその隣接分野で使われていたメタデータスキーマを比較検討。
- ▶ その結果、SPASEを採用することに決めた。
- ▶ SPASEの思想とIUGONETの思想がマッチしたのも大きい。
  - ▶ データ自体の共有には様々な障害があるが、メタデータはそういう縛りは無いはず、だからメタデータ共有から始めよう by Shing Fung @GSFC, NASA

# SPASEベースのIUGONET共通メタデータフォーマットで

**SPASE** Space Physics Archive Search and Extract

<https://spase-group.org/>

- ▶ STP分野、MTI分野のデータを記述できる要素・単語が揃っている。
- ▶ オープンソースとして公開、ドキュメントやツールが整備されている。
- ▶ VxOで使用されるなどde facto standardの有力候補。
- ▶ 可視化属性、データアクセス権に関する情報も格納可能。
- ▶ XMLファイルなのでその気になれば人間も読める。

## ▶ IUGONETで改良した点

- ▶ 観測がカバーする緯度・経度範囲を記述する要素を追加。
- ▶ 太陽撮像データで使用される座標系の定義を追加。
- ▶ DOI付与に利用、ORCIDも?
- ▶ シミュレーション拡張に寄与



VxO懇親会の様子  
(@ AGU Fall 2009)

# 蓋を開けてみたら「データモデル」だった これらだけでデータの世界を記述できる!

## SPASEデータモデルに基づくメタデータ一式

### メタデータ付きデータファイル

ヘッダー部にメタデータ

数値データ部  
0000000000000000  
0000000000  
00000000000000000000  
0000000000000000  
00000000000000000000  
00000000000000000000  
00000000000000000000  
00000000000000000000  
00000000000000000000

あくまでデータファイルの  
付属情報

### データセットを記述

```
<?xml version="1.0" encoding="UTF-8"?>  
<NumericalData>  
... 数値データ XML
```

```
<?xml version="1.0" encoding="UTF-8"?>  
<DisplayData>  
... プロット・動画 XML
```

```
<?xml version="1.0" encoding="UTF-8"?>  
<Catalog>  
... イベントリスト XML
```

### 地上観測サイト・観測衛星

```
<?xml version="1.0" encoding="UTF-8"?>  
<Observatory>  
... XML
```

### 観測機器

```
<?xml version="1.0" encoding="UTF-8"?>  
<Instrument>  
... XML
```

### 研究者など人的リソース

```
<?xml version="1.0" encoding="UTF-8"?>  
<Person>  
... XML
```

- メタデータだけでファイル
- 各リソースで別々、リンク

# SPASEのその後の発展 (自動的にIUGONETも享受)

[https://www.iswat-cospar.org/sites/default/files/2019-08/PSW\\_2018\\_Metadata\\_Resolution\\_V3.pdf](https://www.iswat-cospar.org/sites/default/files/2019-08/PSW_2018_Metadata_Resolution_V3.pdf)

- ▶ NASAがfundした太陽圏衛星ミッションはSPASE形式のメタデータを提供することが義務に。
- ▶ COSPARがSPASEをmetadata standardとしてrecommended (2018)。
- ▶ シミュレーション拡張 → SPASE本体へ

**COSPAR Panel on Space Weather (PSW) Resolution on Metadata Standards.**  
*Accepted at COSPAR PSW Business Meeting on July 18, 2018 (revised December 6, 2018)*

Taking into account that:

1. It is in the general interest of the international space physics and space weather community that data be made as widely accessible as possible.
2. The open exchange of data benefits from accurate, complete, and standardized structured metadata.
3. The ILWS-COSPAR Roadmap recommends to standardize metadata and harmonize access to data and model archives.
4. There are a variety of metadata standards, few of which meet the full needs of the community to document and enable discovery of available diverse data.
5. The SPASE (Space Physics Archive Search and Extract) metadata standard has demonstrated that it is comprehensive and mature and can meet the needs of the community,
6. That sufficiently detailed metadata enhances data access and discovery by applications and through services like HAPI (Heliophysics Application

Therefore, it is recommended that:

1. SPASE be the metadata standard to describe space physics and space weather data.
2. Funding agencies provide encouragement and adequate support to enable data produced by projects to be described with SPASE metadata.

# 今思えば...的な雑感

- ▶ 総じて、SPASEを選んでおいて良かった... (本音)
  - ▶ 今ほどの完成度や適用・派生状況は無かったが、限られた情報と熟慮で選択
  - ▶ 個人的には当時学んだり考えたことや人脈が今の仕事につながっている。
- ▶ とは言え、大量のXMLファイルとの格闘というレガシーを残す
  - ▶ 処理が重い、ツールもJavaばかりで研究者にはとっつきにくい
  - ▶ 今なら生成AI等で下書きできないだろうか。
  - ▶ 怪我の功名: XMLアレルギーは無くなった → 現在のPDS仕事につながった
- ▶ 限られた不十分な情報から考え抜いて最適解を選び出す経験
  - ▶ 十分な情報があれば最適解を得るのはそれ程難しくない。でもほとんどの場合、不十分な情報・知識・経験しかない状態で選択を強いられる。
  - ▶ 研究者の手弁当の技術と最新ITとの混合率で悩む。
  - ▶ その後のERG-SCでも度々選択の場面、現在のBepiColombo/MMO衛星データ開発でも



荻野先生  
(名大ISEE 元教授)  
STEL, ISEEでの  
IUGONETの取りまとめ  
(~2013)  
九工大工学部出身!

# IUGONETメタデータの行く末

シミュレーションデータにもメタデータを振って何かの検索システムから到達可能にしたい。

- ▶ SPASEにシミュレーション拡張がなされ、特にIUGONETがGAIAモデル等のメタデータ登録を通じて色々なアイデアやユースケースをインプットしている。しかしまだまだ単語の拡充や定義の精査、新しいカテゴリー(観測でのInstrumentとか?)の実装が必要なように見える(素人感)。
- ▶ ただしデータをアーカイブするDBがあって、それにメタデータ及びメタデータDBを付随させるのが正統なやり方。
  - ▶ 荒唐無稽かもだが、ジオスペースの実証型モデルのみから始めてもいいので、シミュレーションデータをアーカイブするデータセンターみたいなものを作れないか？ IUGONET型の複数拠点を束ねたバーチャル拠点にしてもいい。
  - ▶ 地球超高層大気モデルの結果は他分野にもニーズがあると想像。

# IUGONETメタデータの行く末

---

メタデータXMLファイルを作る労力をもっと削減する必要。  
またはデータキュレーターのようなキャリアパスを確立してその人らに任せる。

- ▶ 数百行レベルのXMLファイルの作成はなかなか敷居が高い。
  - ▶ 惑星科学分野の標準データアーカイブ形態であるPlanetary Data System (PDS)に至っては、もはや研究者が片手間でできるレベルを超えている。
  - ▶ ChatGPT or 専用に開発したMLで5-6割下書きして残りは人間がrefineする、みたいなワークフローは可能？
- ▶ キュレーターが分野研究者出身の人だと色々都合が良い。
  - ▶ 専門知識や、分野の経緯のような無形な情報も知っている。
  - ▶ IT専門家主導 or 丸投げは、中期的なメンテや将来の引き継ぎの観点で破綻しやすい。

# IUGONETメタデータの行く末

---

## (超)長期保存の策を練らねばならぬ

- ▶ 地球物理観測データは100年スケールで保存して将来の研究に耐えうる状態を保っておく必要がある。
  - ▶ デジタルで体系化しただけでもその可能性は広がった。
- ▶ データモデルの維持・更新
  - ▶ 何となくSPASEはPDSとマージしていくような気がする。
- ▶ その人的リソースも
  - ▶ IUGONETも15周年。初期開発メンバーも若くないし、同様にSPASEの中の人たちも結構なお歳

## まとめと結論

---

祝IUGONET15周年! この機会に、IUGONETメタデータというシステムがどのように生まれ発展してきたかを概観した。

- ▶ 総じてSPASEベースのメタデータ形式をIUGONETで採用して良かった。
  - ▶ SpatialCoverage, シミュレーション拡張、DOIなど、IUGONETからの寄与もあった。
- ▶ 大量のXMLを編集・維持するためのリソースを確保し続ける必要。
- ▶ 超高層物理、太陽圏科学の「観測データ」に関するメタデータのフレームワークはできた。次のフロンティアは(まずは実証型)シミュレーションのデータか。