

2016.10.18 IUGONET研究集会@極地研2F大会議室



図書館員による IUGONET メ タデータ作成実験経過報告

情報・システム研究機構 国立極地研究所
南山 泰之

minamiyama@nipr.ac.jp

ORCID ID:0000-0002-7280-3342

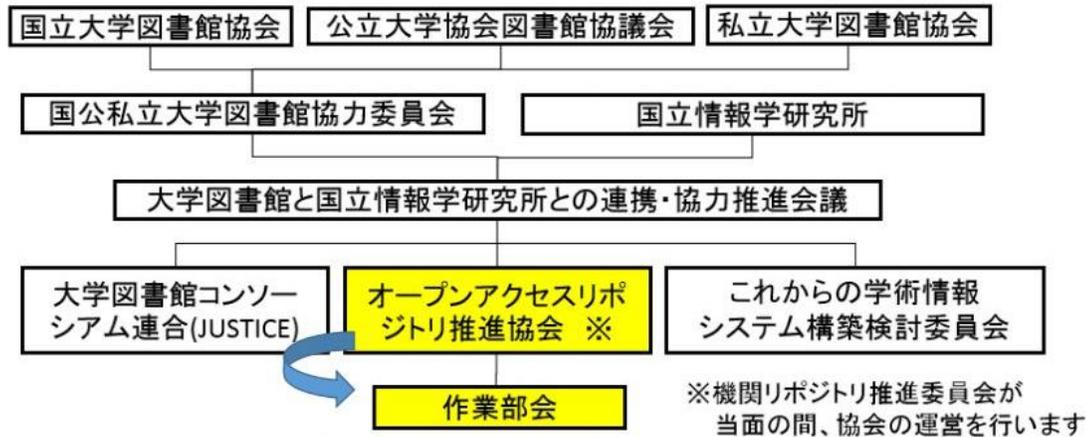
実験の概要

- ▶ 図書館職員によるIUGONETメタデータ作成を通じて、
 1. 研究者へのインセンティブ
 - ▶ 研究者、特にメタデータ作成業務従事者の負担軽減
 - ▶ データベースの持続可能性向上？
 - ▶ 図書館が持つ流通面での知見の提供
 2. 図書館による新規サービス展開
 - ▶ 研究者との協働による、より直接的な研究支援

の可能性を探る。

図書館ができること

ネットワークの活用



参加機関は、運営委員会や作業部会に委員を派遣することにより、事業に参画

<http://www.nii.ac.jp/content/cpc/>

メタデータ運用

アイテムタイプ選択 マッピング設定 確認

アイテムタイプ: sample metadata 未設定

メタデータ名	入力タイプ	Dublin Core	junii2	言語
タイトル	テキスト	title	title/alternative	日本語
タイトル(英)	テキスト	title	title/alt	英語
言語	選択式(プルダウン)	language	language	指定なし
公開日	日付	date	date	指定なし
キーワード	テキスト	subject	subject	日本語
書名	氏名	未設定	未設定	指定なし
抄録	テキストエリア	description	description	指定なし
関連サイト	リンク	未設定	未設定	指定なし
書籍情報	雑誌名(巻号)	未設定	未設定	指定なし
	巻		jtitle	日本語
	号		jtitle	英語
	開始ページ		volume	指定なし
	終了ページ		issue	指定なし
	発行年		spage	指定なし
			epage	指定なし
			dateofissued	指定なし

① Dublin Coreの語彙を選択する

② JuNii2を選択する

③ JuNii2の語彙を選択する

④ Learning Object Metadata, LIDO, SPASEもプルダウンで切り替えて語彙を設定する

「書籍情報」属性の語彙は固定されている

junii2
junii2
Learning Object Metadata
LIDO
SPASE

< 戻る 次へ >

https://test.repo.nii.ac.jp/weko/help/ja/html/itemtype_mapping.html#mapping

連携によって解決できそうなこと



Current issues in the upper atmospheric research in Japan ...

- Database has been **maintained individually by each university/institute**, so it is difficult for researchers to discover and access the data due to lack of information of them.
- Database has been built and maintained by **domain researchers**.
- Due to **a variety of data**, collection of the data and metadata is time consuming.
- File format is different for each instrument type, thus it usually takes time to analyze many kinds of data.



図書館のネットワーク活用



メタデータ運用に特化した人材

なぜIUGONET？

- ▶ 分野横断のデータベース構築
 - ▶ “本プロジェクトでは、超高層大気に関連する地上観測データのメタデータ・データベースを構築し、これまで個別の研究機関で分散管理されてきた観測データの流通を図ります”
 - ▶ 図書館の基本的な機能と親和性が高い
- ▶ 明確かつ安定したメタデータ標準
 - ▶ “国際的な標準化の流れの中で超高層分野のデータにもっともマッチし、かつ汎用性・拡張性がある「SPASE」を、IUGONET共通メタデータフォーマットのベースに採用することに決定しました”
 - ▶ 運用レベルでメタデータを扱うことが可能
- ▶ 持続可能性
 - ▶ 運用実績、開発体制など

<http://www.iugonet.org/index.html>

パイロット (6月～)

図書職員とのコラボレーション

No.	難易度	内容	所感 (図書)	所感 (IUG)
1	低	既存のメタデータ(XML)を機械的に修正する	◎ 容易	◎ OK
2	中-	既存のメタデータ(XML)を新しいスキーマバージョンに則したものに変更する	○ 出来た	○ OK
3	中	PIにインタビューしながら新しくサイエンスメタデータを作成する	進行中	○ OK
4	中	外部サーバを使って単純なデータ公開・更新を行う	(予定)	
5	中+	自前サーバを使って単純なデータ公開・更新を行う	(予定)	
6	高	IUGONETのようなサービスのデータ公開部分を運用する	(予定)	
7	高+	IUGONETのようなサービスのデータ登録部分を運用する		

サイエンス成果の創出に向けたデータアーカイブと解析環境の融合. 梅村宜生ほか
第3回オープンサイエンスデータ推進ワークショップ@京大より抜粋

Metadata Schema (Spase)



Welcome to the SPASE Group

The Space Physics Archive Search and Extract (SPASE) effort is a HelioPhysics community-based project with the goals of:

- Facilitating data search and retrieval across the Space and Solar Physics data environment with a common metadata language
- Defining and maintaining a standard Data Model for Space and Solar Physics interoperability, especially within the HelioPhysics Data Environment
- Using the Data Model to create data set descriptions for all important HelioPhysics data sets
- Providing tools and services to assist SPASE data set description creators as well as the researchers/users
- Working with other groups for other HelioPhysics data management and services coordination as needed

The Space Physics Archive Search and Extract (SPASE) effort is implemented by the SPASE Consortium which is composed of representatives of the international HelioPhysics data community. The SPASE Working Group is currently the only international group supporting global data management for Solar and Space Physics.

Learn more about the SPASE group.

Products

The SPASE Group generates three "products". First is the SPASE [Metadata Model](#) which is an information model for describing the elements of the helioPhysics data environment. Second is a set of resources and protocols to enable the exchange of information. Third are tools for developing and validating resource descriptions.

A complete set of documentation is also available.
Tutorials and instructions are available at the [SPASE School](#).

Data Model Document

Current Version (2.2.0)
Released: 2011-01-06
All documents
History of changes

Schema

Current Release (2.2.0)
updated: 2011-01-05
View all versions

Tools

Data Dictionary Search
Data Model Tree
Data Model Explorer

Registry Server
Resource ID maker
Editor (web)
XML Validator

Services

SMWG Registry
Naming Authority List

Looking for:

Vintage web site

SPASE

<http://www.spase-group.org/>

NASA, 米国の大学・研究機関及びヨーロッパの関連機関からの研究者から成るコンソーシアムで策定・随時update

基本的に太陽、惑星間空間、地球磁気圏の人工衛星観測データを念頭に、それらに関連する研究リソースを包括的に表現するデータモデルに基づいて作られた、メタデータフォーマット

```
<?xml version="1.0" encoding="UTF-8"?>
<!-- Automatically created based on the dictionary stored at http://www.spase-group.org -->
<!-- Version: 2.2.6 -->
<!-- Generated: 2015-09-09 09:01:12.550 -->
<xsd:schema
  targetNamespace="http://www.spase-group.org/data/schema"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns:vc="http://www.w3.org/2007/XMLSchema-versioning"
  xmlns:spase="http://www.spase-group.org/data/schema"
  elementFormDefault="qualified"
  attributeFormDefault="unqualified"
  vc:minVersion="1.1"
  version="2.2.6"
>
  <xsd:element name="Spase" type="spase:Spase" />
  <xsd:complexType name="Spase">
    <xsd:annotation>
      <xsd:documentation xml:lang="en">
        Space Physics Archive Search and Extract (SPASE).
        The outermost container or envelope for SPASE
        metadata. This indicates the start of the
        SPASE metadata.
      </xsd:documentation>
    </xsd:annotation>
    <xsd:sequence>
      <xsd:element name="Version" type="spase:Version" minOccurs="1" maxOccurs="1" />
      <xsd:choice minOccurs="1" maxOccurs="unbounded">
        <xsd:element name="Catalog" type="spase:Catalog" />
        <xsd:element name="DisplayData" type="spase:DisplayData" />
        <xsd:element name="NumericalData" type="spase:NumericalData" />
        <xsd:element name="Document" type="spase:Document" />
        <xsd:element name="Granule" type="spase:Granule" />
        <xsd:element name="Instrument" type="spase:Instrument" />
        <xsd:element name="Observatory" type="spase:Observatory" />
        <xsd:element name="Person" type="spase:Person" />
        <xsd:element name="Registry" type="spase:Registry" />
        <xsd:element name="Repository" type="spase:Repository" />
        <xsd:element name="Service" type="spase:Service" />
        <xsd:element name="Annotation" type="spase:Annotation" />
      </xsd:choice>
    </xsd:sequence>
    <xsd:attribute name="lang" type="xsd:string" default="en"/>
  </xsd:complexType>
  <xsd:complexType name="Catalog">
    <xsd:annotation>
      <xsd:documentation xml:lang="en">
        A tabular listing of events or observational
        notes, especially those that have utility
        in aiding a user in locating data. Catalogs
        include lists of events, files in a product,
        and data availability. A Catalog resource
        is a type of "data product" which is a set
        of data that is uniformly processed and formatted,
        from one or more instruments, typically spanning
        the full duration of the observations of the
        relevant instrument(s). A data product may
        consist of a collection of granules of successive
        time spans, but may be a single high-level
        entity.
      </xsd:documentation>
    </xsd:annotation>
    <xsd:sequence>
      <xsd:element name="ResourceID" type="spase:ResourceID" minOccurs="1" maxOccurs="1" />
      <xsd:element name="ResourceHeader" type="spase:ResourceHeader" minOccurs="1" maxOccurs="1" />
    </xsd:sequence>
  </xsd:complexType>

```

<http://www.iugonet.org/mdformat.html>

メタデータの作成・編集

- ① Validatorによるエラーチェック
- ② 担当者へメールによる確認依頼
- ③ メタデータ登録

```
133.57.20.53:22 - minamiyama@iugonet0:~ VT
ファイル(F) 編集(E) 設定(S) コントロール(O) ウィンドウ(W) ヘルプ(H)
###INFO> XMLFile = [/home/minamiyama/NIPR/Repository/NIPR/WDC_Aurora_NIPR.xml]
#RESULT> Parse...[OK]
#RESULT> Validation...[OK], SCHEMA = [/home/minamiyama/md_validator/xsd/iugonet-2_2_6_1.xsd]
-----
###INFO> XMLFile = [/home/minamiyama/NIPR/Repository/NIPR/IUGONET_DB_NIPR.xml]
#RESULT> Parse...[OK]
#RESULT> Validation...[OK], SCHEMA = [/home/minamiyama/md_validator/xsd/iugonet-2_2_6_1.xsd]
-----
###INFO> XMLFile = [/home/minamiyama/NIPR/Repository/NIPR/CDC_Web_NIPR.xml]
#RESULT> Parse...[OK]
#RESULT> Validation...[OK], SCHEMA = [/home/minamiyama/md_validator/xsd/iugonet-2_2_6_1.xsd]
-----
###INFO> XMLFile = [/home/minamiyama/NIPR/Repository/NIPR/EISCAT_DB_NIPR.xml]
#RESULT> Parse...[OK]
#RESULT> Validation...[OK], SCHEMA = [/home/minamiyama/md_validator/xsd/iugonet-2_2_6_1.xsd]
-----
###INFO> XMLFile = [/home/minamiyama/NIPR/Repository/NIPR/NWI_MOVIE_WEB_NIPR.xml]
#RESULT> Parse...[OK]
#RESULT> Validation...[OK], SCHEMA = [/home/minamiyama/md_validator/xsd/iugonet-2_2_6_1.xsd]
-----
###INFO> XMLFile = [/home/minamiyama/NIPR/Repository/NIPR/ASI_SPA_WEB_NIPR.xml]
#RESULT> Parse...[OK]
#RESULT> Validation...[OK], SCHEMA = [/home/minamiyama/md_validator/xsd/iugonet-2_2_6_1.xsd]
-----
[RESULT] -----
#RESULT> TOTAL = [193], (NORMAL = [193], ERROR = [0])
#RESULT> COMPLETE!!
-----
[minamiyama@iugonet0 ~]$
```

メタデータの作成・編集

Inter-university Upper atmosphere Global Observation NETwork

IUGONET
Metadata DB for Upper Atmosphere

UDAS web Available!

IUGONET DataSet

Instrument Type	Observed Region
Satellite: <input type="checkbox"/> AKEBONO <input type="checkbox"/> CHAMP	
Ground-Based: <input type="checkbox"/> SMART (Telescope) <input type="checkbox"/> DST (Telescope) <input type="checkbox"/> FMT (Telescope) <input type="checkbox"/> Refractor (Telescope) <input type="checkbox"/> Geomagnetic Indices <input type="checkbox"/> WDC Geomag., Kyoto <input type="checkbox"/> Magnetometer (KMO) <input type="checkbox"/> Magnetometer <input type="checkbox"/> MAGDAS <input type="checkbox"/> MM210 <input type="checkbox"/> AWS <input checked="" type="checkbox"/> All Sky Imager <input type="checkbox"/> MU Radar <input type="checkbox"/> MF Radar <input type="checkbox"/> MW Radar <input type="checkbox"/> X-Band Radar <input type="checkbox"/> Na-Lidar <input type="checkbox"/> EISCAT <input type="checkbox"/> OMTI <input type="checkbox"/> SuperDARN <input type="checkbox"/> VLF/ELF <input type="checkbox"/> Ionosonde <input type="checkbox"/> Radiosonde <input type="checkbox"/> BL/LT/WP Radar <input type="checkbox"/> Others	
Keyword: <input type="text" value="keogram data"/>	
Timespan: <input type="text"/> To <input type="text"/> Set Detail	
<input type="button" value="Search"/>	

Search Results:

[Text](#) [Plot](#) : Contains Summary Plot : Create Pl

Ground-Based

All Sky Imager

- [Plot/Movie Data](#) [All-sky auroral image taken by the white-light all-sky camera at Husafell, Iceland.](#)
- [Plot/Movie Data](#) [All-sky auroral image taken by the white-light all-sky camera at Syowa Station, Antar](#)
- [Plot/Movie Data](#) [Keogram data derived from auroral images taken by the white-light all-sky imager at](#)
- [Plot/Movie Data](#) [Keogram data derived from auroral images taken by the white-light narrow-field-of-vi](#)

All Rights Reserved.
Copyright (c) 2009-2016 IUGONET

[List](#) > [Instrument Type](#) > [Ground-Based](#) > [All Sky Imager](#)

Plot/Movie Data Keogram data derived from auroral images taken by the white-light all-sky imager at Tromso

Description:

Keogram data derived from auroral images taken by the white-light all-sky imager (body: WATEC WAT-120N+, lens: FUJINON YV2.2×1.4A-SA2) at Tromso, Norway. The all images are originally recorded in the JPEG format with a temporal resolution of ~0.25s.

Acknowledgement: The data in this web site are the intellectual property of National Institute of Polar Research, JAPAN. They may be freely used for the purpose of illustration for teaching and for non-commercial scientific research, provided that the source is acknowledged. Substantial use of the data should be discussed at an early stage with the PI of each instrument. NIPR researchers are hoping collaborative studies and fruitful scientific outputs using these data.

ReleaseDate: 2016-10-01T12:00:00

Contact (PrincipalInvestigator):

Yasunobu Ogawa, National Institute of Polar Research, Japan, yogawa [at] nipr.ac.jp

Contact (CoInvestigator):

Akira Kadokura, National Institute of Polar Research, Japan, kadokura [at] nipr.ac.jp

Contact (CoInvestigator):

Yoshimasa Tanaka, National Institute of Polar Research, Japan, ytanaka [at] nipr.ac.jp

Contact (CoInvestigator):

Tetsuo Motoba, National Institute of Polar Research, Japan, motoba.tetsuo [at] nipr.ac.jp

Contact (CoInvestigator):

Keisuke Hosokawa, The University of Electro-Communications, keisuke.hosokawa [at] uec.ac.jp

Contact (MetadataContact):

Yasuyuki Minamiyama, National Institute of Polar Research, Japan, minamiyama [at] nipr.ac.jp

AccessInformation:

URL: http://133.57.20.115/www/optical/watec/wmi_all/schedule_calendar.php

Availability: Online

Access Rights: Open

Format: PNG

Measurement Type: ImageIntensity

Time Span:

StartDate: 2010-11-09T00:00:00Z

StopDate: -P7D

Observed Region: Earth.NearSurface.Ionosphere.ERegion

Observed Region: Earth.NearSurface.Ionosphere.FRegion

ここまでの検証

- ▶ 既存の情報源をメタデータ化する作業は、図書館の
目録作成業務の延長
 - ▶ 標準化されたスキーマや入力ルールが定められていれば、
分野を問わず対応可能
 - ▶ 連携により、お互いの強みを活かした形で作業を効率化できる
 - ▶ もっとも、データ取得者による確認作業は省略できない

- ▶ 人的な持続可能性の問題
 - ▶ 作業は慣れや労力を要するが、**マニュアル作成によって**
標準化・引き継ぎが可能
 - ▶ メタデータ作成に要する時間は、図書や地図とあまり変わらない
 - ▶ 図書館のネットワークで共有することで、全体の知見になる

コード	一般資料種別	コード	特定資料種別
なし	下記のいずれでもないもの	なし l t	下記のいずれでもないもの 大活字本 (largeprint) 触知資料 (tactil)
a	地図資料 (cartographic material)	a b c g j k q r s y z	天体儀 (celestial globe) 惑星儀 / 月球儀 (planetary or lunar globe) 地球儀 (terrestrial globe) ダイアグラム (diagram) 地図 (map) 側面図 (profile) 模型 (model) リモート・センシング画像 (remote-sensing image) 断面図 (section) 景観図 (view) 上記のいずれでもないもの
b	文字資料 (点字) (brail)	なし	
c	楽譜 (印刷) (printed music)	a b c d e g h m u z	スコア (full score) ミニチュアスコア (full score, miniature or study size) 鍵盤楽器伴奏譜 (accompaniment reduced for keyboard) ヴォイススコア (voice score) コンデンススコア / ピアノコンダクタースコア (condensed score or piano-conductor score) クローズスコア (close score) 邦楽譜 複合形態の楽譜 (multiple score format) 不明 (unknown) 上記のいずれでもないもの (other than score format)
d	文字資料 (書写資料) (manuscript text)		GMD「なし」に対応するSMDのうち適切なコードを用いる
e	地図 (書写資料) (manuscript map)		GMD「a」に対応するSMDのうち適切なコードを用いる
f	楽譜 (書写資料) (manuscript music)		GMD「c」に対応するSMDのうち適切なコードを用いる
g	静止画像 (投影) (projected graphic)	c d f	フィルムストリップ (カートリッジ) (filmstrip cartridge) フィルムストリップ (filmstrip) フィルムストリップ (その他) (other filmstrip type)

今後やりたいこと

- ▶ 運用レベルでの支援
 - ▶ メタデータの検証、データクリーニング
 - ▶ エンリッチメント?
- ▶ 水平展開
 - ▶ 複数人による安定的なメタデータ作成サービス提供
 - ▶ できれば組織横断で
- ▶ 図書館側の知見の提供
 - ▶ 機関リポジトリとの連携
 - ▶ メタデータのハーベストによる流通支援
 - ▶ ライセンス、識別子、etc...