

WMO 情報システム (WIS) における
メタデータ技術

豊田英司

toyoda.eizi@gmail.com

2011-05-25

於 地球惑星科学連合大会

今日のお話

- メタデータ形式
- 利用者インターフェイス
- 機関を超えたメタデータ交換
- 今後の課題

そもそも WIS とは

- WIS = WMO Information System
 - WMO = 世界気象機関
- 従来からの GTS (全球気象通信網) を継承
 - 現業気象業務のリアルタイム情報交換
 - インターネット以前からのレガシーシステム
- すべての WMO ・ 関連国際計画の基盤に進化
 - 当面はカタログ (メタデータ) 作成に注力

Organisational structure

- GISC: global information system centre
 - information catalogue of entire WIS
 - global distribution information on web
 - network management
- DCPC: data collection & product centre
 - regional or programme-wide activities
- NC: national centre
 - national activities

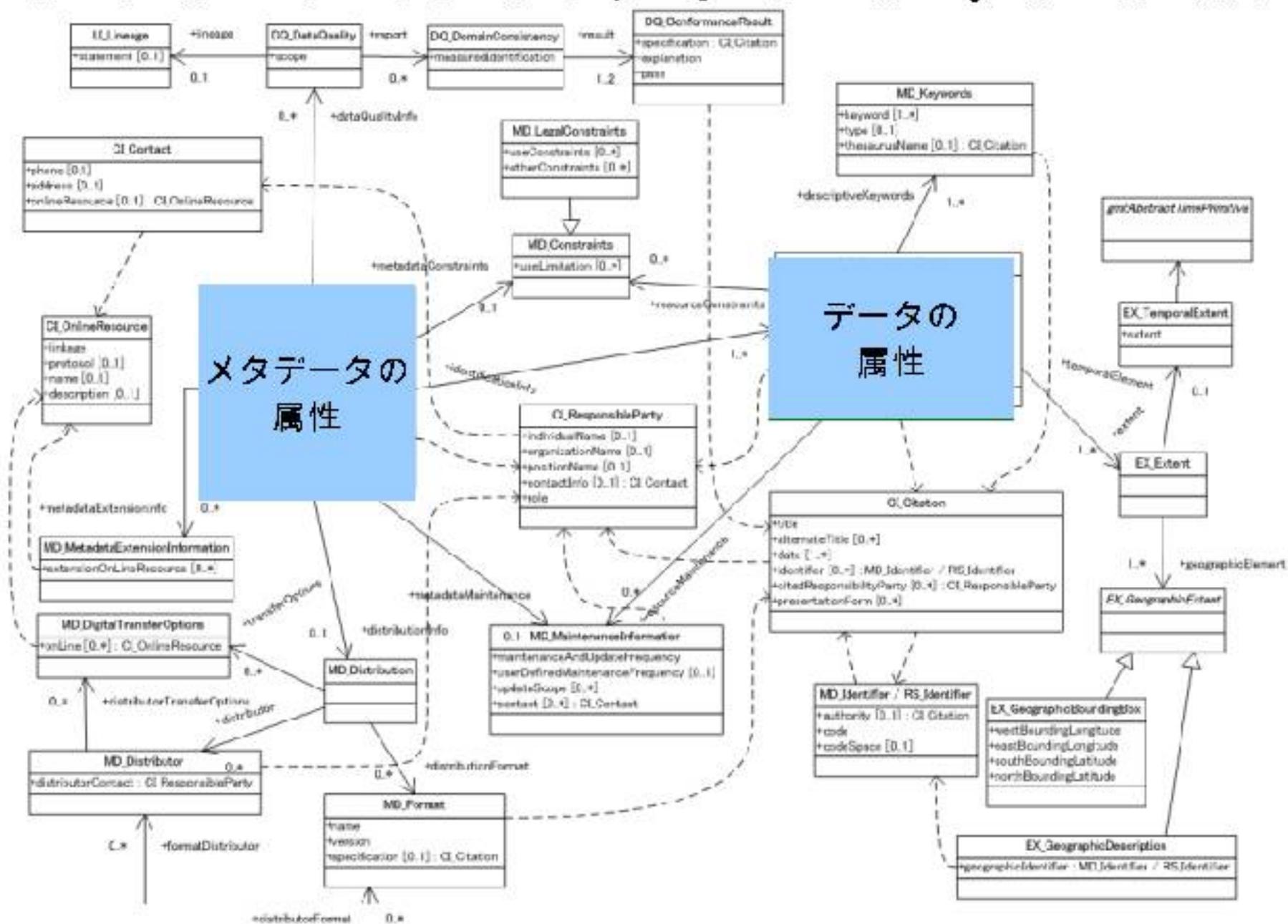


要するに
主要国では全 WMO の
データカタログを作る

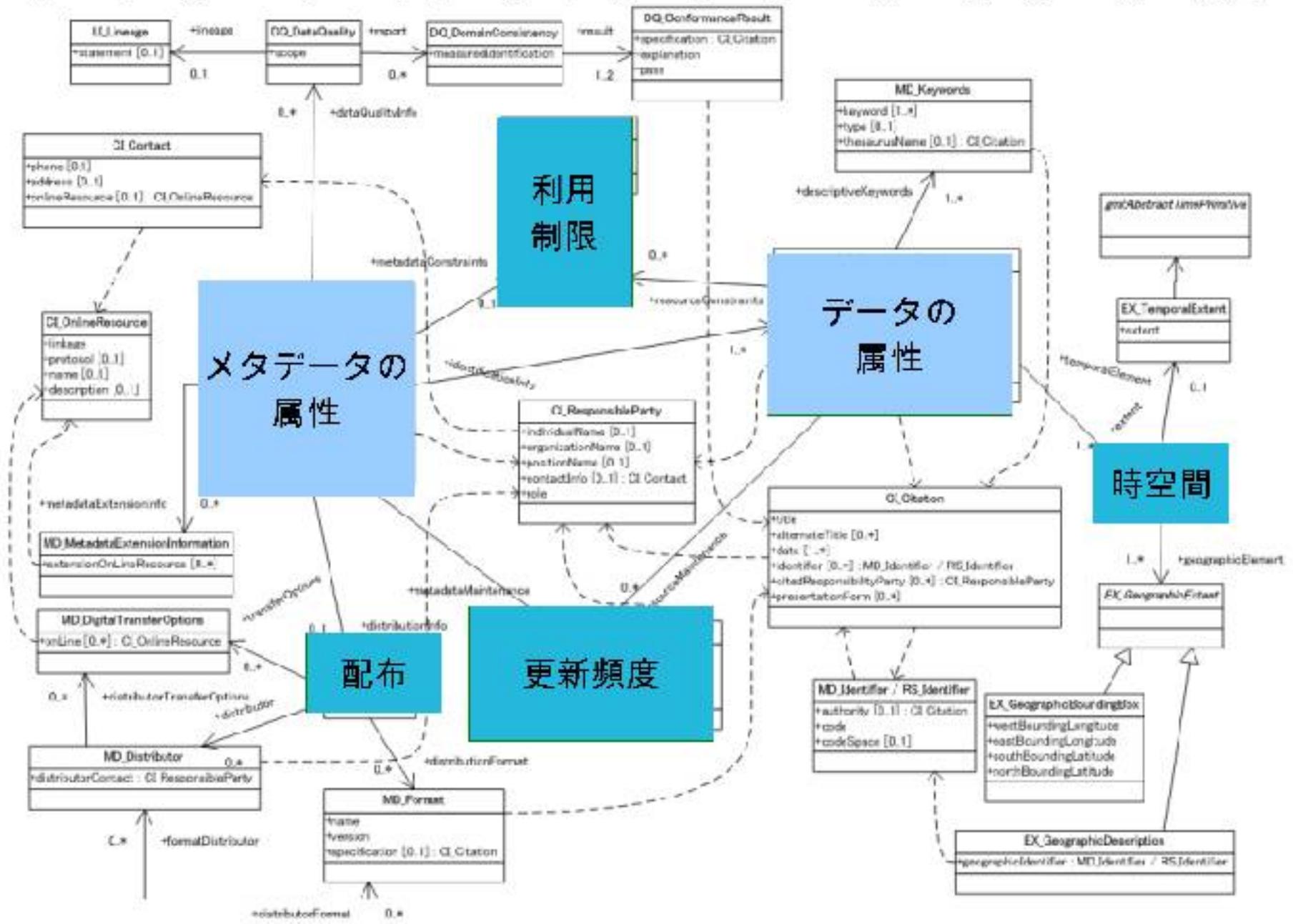
ついにメタデータ構造の議論始まる

- WMO は ISO 19115 を採用 (2003)
- 当初の議論は ISO の拡張
 - より完全な記述を目指して ...
- WMO Core Metadata Profile v1.1 (2009)
 - 拡張要素は全廃
 - 他者が理解できないし標準スキーマで検証できない
- 今の議論は、さらに絞込み
 - 誰でもが記入すべき中核は何か

ここまでコンパクトになりました



ここまでコンパクトになりました



スキーマ規模

- WMO プロファイルは小さくなったとはいえ、単純型まで展開すると 280 要素（現状案）
 - 誰が全部覚えられるでしょうか
 - でも必須項目は 14 要素
- 現実にワークするセットは 30 要素程度が限界
 - レガシーなカタログ類も 10 ~ 30 程度 例

入力について

- ウェブ・オンライン編集
- エクセル入力
 - JAMSTEC CrDAP のは VBA で XML 作成・検証までやっていますすごい
 - でもうちはマクロをやめようかと思っています
 - CSV 等プログラム作成しやすい形式にして、ルーチンの提供体制を確立したい

利用者インターフェイス

SRU プロトコル

- 図書館情報方面で開発
 - ISO 23950 = ANSI Z39.50 の代替を目指す
- 技術的特徴は OGC WMS 等と類似
 - 問い合わせは CGI パラメタ (又は SOAP)
 - 結果はメタデータ記録が複数入った XML

SRU とマッシュアップ

- マッシュアップ：ウェブサービスをサブルーチンのように使って第三者が高付加価値サイトを作ること
 - 巷の twitter API とか Google 検索窓とか
- SRU の結果は XML なので第三者サイトでの利用が容易
 - ただし JavaScript を使う際は XSS 問題に注意

SRU のインデックス

- 「 title = "synoptic" 」 の title とか
- 並列検索するなら標準化が必要
- WMO ではようやく 19115 要素 → SRU インデックスの対応の議論が始まった

WMO で議論されている SRU インデックス

文字型

- author
- title
- abstract
- identifier
- keyword
- type
- crs

時刻型

- creationDate
- modificationDate
- publicationDate
- beginningDate
- endingDate

経緯度範囲型

- bounds

基本的に
OGC CSW 用の
19115 マッピングを
基にしている

WMO ホームページの検索画面

<p>Text terms <input type="checkbox"/> include</p>	<p>Find: <input type="text" value="hack"/> in: <input type="text" value="FullText"/> <input type="text" value="AND"/></p> <p>Find: <input type="text"/> in: <input type="text" value="FullText"/> <input type="text" value="AND"/></p> <p>Find: <input type="text"/> in: <input type="text" value="FullText"/></p>	<p>皆さんの フォームとの違いは 日付の種類があること くらい</p>
<p>Find metadata records by date <input type="checkbox"/> include</p>	<p>AND using: a start date of: <input type="text" value="1"/> <input type="text" value="Jan"/> <input type="text"/> used against <input type="text" value="CreationDate"/></p> <p>AND using: an end date of: <input type="text" value="1"/> <input type="text" value="Jan"/> <input type="text"/> used against <input type="text" value="CreationDate"/></p>	
<p>Keyword Search</p>	<p>AND using: WMO subject term: <input type="text" value="none"/></p>	
<p>Spacial Search <input type="checkbox"/> include</p>	<p>AND: includes the following coordinates:</p> <p>North: <input type="text"/></p> <p>West: <input type="text"/> East: <input type="text"/></p> <p>South: <input type="text"/></p>	



機関を超えたメタデータ共有

OAI-PMH プロトコル

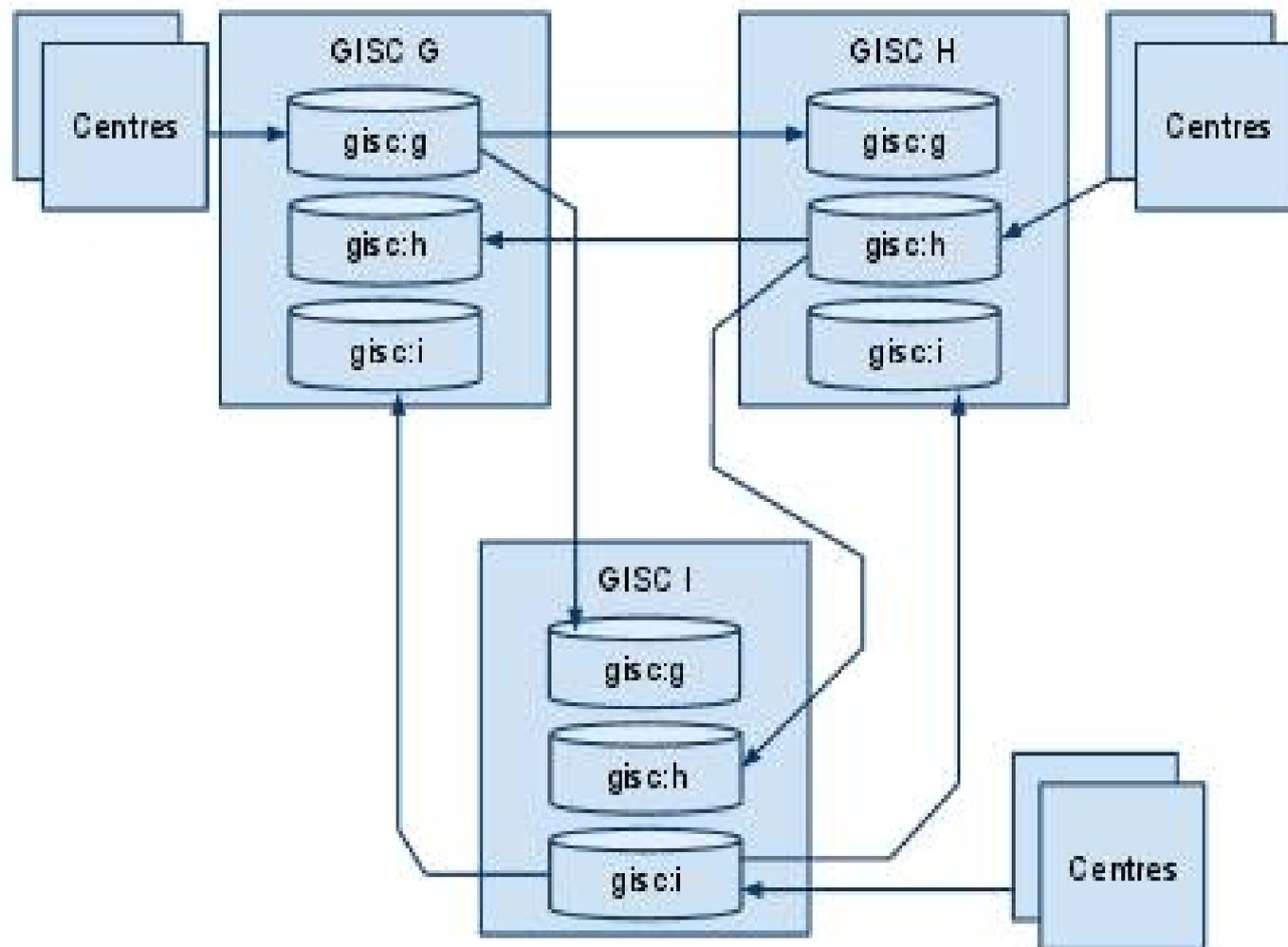
- 図書館情報方面で開発
 - RSS 的更新通知
 - 時間範囲指定 (抜け防止)
- 技術的特徴は SRU と類似
 - 問い合わせは CGI パラメタ
 - 結果はメタデータ記録が複数入った XML

巡回トポロジでの OAI-PMH

- WIS のネットワークには唯一の中心がない
- 元々 OAI-PMH は非巡回ネットワークが前提
 - 同じ identifier を更新することができる
 - 識別子と日付で同定、といたいところだが
 - 日付は現サイトの更新日、出所日付と無関係
 - これは取りこぼし防止のため必要な規定
- 循環トポロジでは更新が無限ループする

OAI-PMH セットによる非巡回化

- 発信センターのレコードだけを取得する



WIS での相互運用試験

- いくつかの実装が利用されている
 - jOAI ベース (複数)
 - Geonetwork
 - スクラッチアップ
- 稼動監視→相互運用性問題を発見、解決
 - GML 名前空間互換性問題
 - status="deleted"

GML 名前空間互換性問題

- ISO 19139 地理メタデータ標準は、一部データ構造の表現に GML (ISO 19136) を利用
- 条文では ISO 19136 を参照
 - xmlns:gml="http://www.opengis.net/gml/3.2"
- 2007 年配布の XML スキーマでは ISO 19136 のドラフト を利用
 - xmlns:gml="http://www.opengis.net/gml"
- 前者で統一、オンラインフィルタで解決

今後の課題

メタデータ作成維持は大変

- ほうっておくと更新されなくなってしまう
- 既存のカタログ作成活動を、利用・継承・改善

例：GTS 電文カタログ

- CSV 形式で流通
 - 全 12 カラム：ただし各列の定義は曖昧
- 利用：現在各国で一方向変換
- 継承：いずれはメタデータが主に？
- 改善：ISO の対応項目を選んで定義を明確化？
 - 自己記述データの場合はかなり自動抽出可能
 - 要人手：タイトル・アブスト・利用制約くらい

ご清聴ありがとうございました

Pre-operational at www.gisc.kishou.go.jp