

地球観測データに対するメタデータ登録システムの構築

絹谷 弘子[†] 生駒 栄司[†] 吉川 正俊^{††} 喜連川 優^{†††} 小池 俊雄^{††††}

[†] 東京大学地球観測データ統融合連携研究機構

^{††} 京都大学大学院情報学研究科

^{†††} 東京大学生産技術研究所

^{††††} 東京大学大学院工学系研究科社会基盤学専攻

E-mail: {[†]kinutani, [†]eikoma, ^{††}kitsure}@tkl.iis.u-tokyo.ac.jp, ^{††}yoshikawa@i.kyoto-u.ac.jp,

^{††††}tkoike@hydra.t.u-tokyo.ac.jp

あらまし 気候変動や温暖化現象が我々の生活に様々な影響を与えるようになってきた。災害防止のためには地球観測データを国境を越えて共有し、活用する必要性がますます高まっている。アジアモンスーン地域を対象とした水循環機構の解明と防災、対象地域の水環境に関する情報の共有をめざし2005年にアジア水循環機構 (Asian Water Cycle Initiative: AWCI) が創設された。現在我々はAWCI参加国が保有する地上観測データの統融合利用の基盤整備を行っているところである。我々は各国が保有する地上観測データを我々が保有する地球観測データ統融合コアシステムに投入し、参加各国が相互に利用できる体制を構築する一環としてデータの送付、データ品質管理、メタデータの登録を行う一連のシステムを構築しているところである。本稿では、我々が構築した地上観測データに関するメタデータ登録システムについて紹介する。

キーワード メタデータ, XML, XForms, ISO19115, E-サイエンス

On Implementation of a Metadata Registration System to the Earth Observation Data

Hiroko KINUTANI[†], Eiji IKOMA[†], Masatoshi YOSHIKAWA^{††}, Masaru KITSUREGAWA^{†††}, and
Toshio KOIKE^{††††}

[†] Earth Observation Data Integration & Fusion Research Initiative, The University of Tokyo

^{††} Graduate School of Informatics, Kyoto University

^{†††} Institute of Industrial Science, The University of Tokyo

^{††††} School of Engineering, The University of Tokyo

E-mail: {[†]kinutani, [†]eikoma, ^{††}kitsure}@tkl.iis.u-tokyo.ac.jp, ^{††}yoshikawa@i.kyoto-u.ac.jp,

^{††††}tkoike@hydra.t.u-tokyo.ac.jp

Abstract The climate change and the global warming have come to have various influences on our life. The necessity for crossing the border, sharing the earth observation data, and use has risen more and more for disaster prevention. Asian Water Cycle Initiative (AWCI) was founded in 2005. The aims of AWCI are to better understanding the mechanism of variability in the Asian water cycle, to improve its predictability, and to interpret the information applicable to various water environments in different countries in Asia, and to help to mitigate water-related disasters and promote the efficient use of water resources. We are establishing a common platform for our data integration and fusion system, especially for the in-situ data which AWCI participants have. We are making the series of systems which subscribe the observation data AWCI participants hold, control data quality and register metadata. And AWCI participants will mutually use those data in the earth observation integrate and fusion core system that we develop. In this paper, we introduce our developing metadata registration system for the in-situ observational data.

Key words metadata, XML, XForms, ISO19115, E-Science

1. はじめに

気候変動や温暖化現象が我々の生活に様々な影響を与えるようになってきた。特に豪雨、洪水や干ばつ、森林火災などが頻発するようになり災害防止のためには地球観測データを国境を越えて共有し、活用する必要性がますます高まっている。リモートセンシング技術によって全世界的に観測されている衛星観測データに比べ、地上観測データは古くから観測されているにもかかわらず、各観測機関あるいは各国の観測網において主に利用されているのが現状である。アジアモンスーン地域を対象とした水循環機構の解明と防災、対象地域の水環境に関する情報の共有をめざし 2005 年にアジア水循環機構 (Asian Water Cycle Initiative : AWCI) が創設された。我々は現在、AWCI 参加国が保有する地上観測データの統融合利用の基盤整備を行っている。AWCI 参加各国には、今までに様々な観測機器により観測されたデータが蓄積されている。参加各国を流れる河川の流域を対象として地上観測ステーションが保有する地球観測データを収集し、品質の高いデータを作ることが求められている。さらに各国が保有する地上観測データを我々が保有する地球観測データ統融合コアシステムに投入し、他の地上観測データ、衛星データや気象予測モデル、さらに参加各国が提供するデータを相互に利用できる体制を構築することで、洪水予防、水資源管理に資する情報を提供することが可能となる。このように AWCI 参加国の保有する地上観測データの価値を高める一環として AWCI 参加国からのデータアップロード、データ品質管理、メタデータ登録を行う一連のシステムを構築しているところである。

本稿では、我々が構築した地上観測データに関するメタデータ登録システムについて紹介する。地上観測データに関するメタデータは、地理空間情報を対象として標準化されているメタデータ ISO 19115:2003 [4] を土台に拡張した独自のメタデータを利用している。このメタデータ登録システムで登録されたデータは、先行研究 [8] で設計した地球観測データに対するメタデータ処理システムにおいてメタデータデータベースに格納されデータの利活用促進のために利用される。本稿の構成は次の通りである。2. で地球観測データとメタデータについて、3. では現状のメタデータ登録システムについて、4. では構築する一連のシステムについて紹介し、5. では XForms を利用したメタデータ登録システム、6. ではまとめと今後の課題について述べる。

2. 地球観測データとメタデータ

異なった観測環境や観測機関からのデータを比較検討するためには、メタデータの内容や形式には一貫性を持たせることが重要である。そのためメタデータ形式や内容を標準化する作業が行われている。例えば、地球観測データに対するメタデータ標準にはアメリカ合衆国連邦地理データ委員会 (FGDC : the Federal Geographic Data Committee) の CSDGM (The Content Standard for Digital Geospatial Metadata) [2] や国際標準化機構 (ISO : The International Organization for Stan-

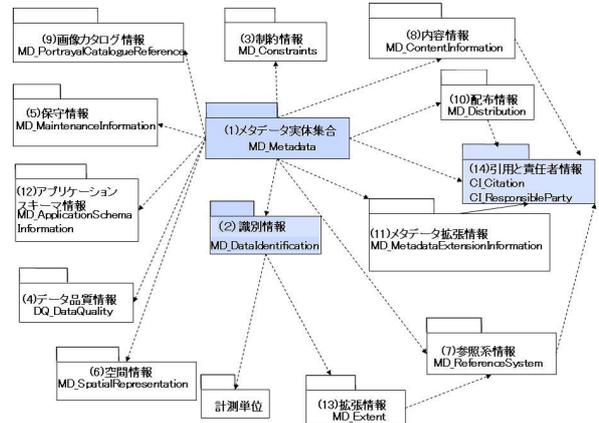


図 1 ISO19115:2003 の構成

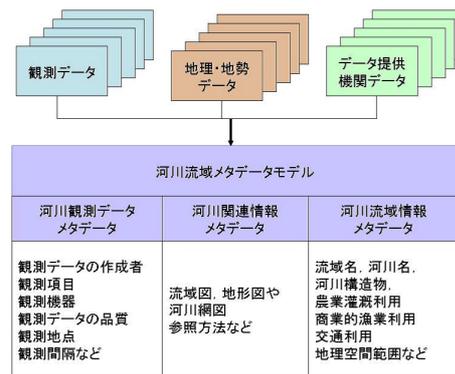


図 2 設計した河川流域を対象としたメタデータ

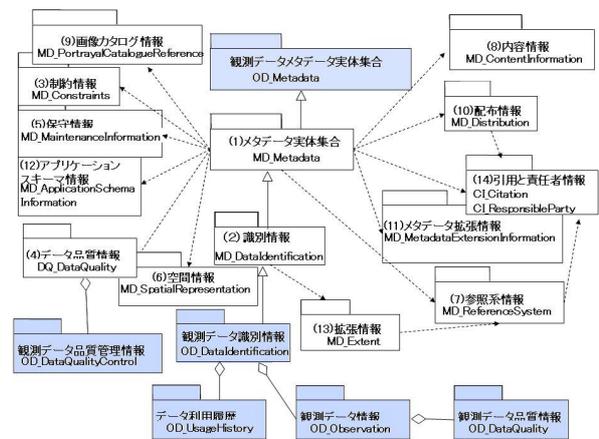


図 3 ISO19115:2003 を拡張した河川観測データメタデータの構成

dards) のメタデータ標準 (ISO 19115:2003) [4] がある。図 1 は ISO19115:2003 を構成するパッケージの関係を表している。必須属性のあるパッケージは (1) メタデータ実体集合 (2) 識別情報と (14) 引用と責任者情報の 3 つであり、その他のパッケージは必要に応じて利用する設計となっている。

AWCI 参加国の地上観測データに対するメタデータ (河川流域メタデータモデル) は ISO 19115:2003 を土台として不足部分を拡張して設計された。設計は、河川流域に関する河川流域情報メタデータ、観測データに関する河川観測データメタデータと河川流域に関連するその他のオブジェクトに対する河川関連情報メタデータの 3 つに分けて行われた。図 2 は各メタデー

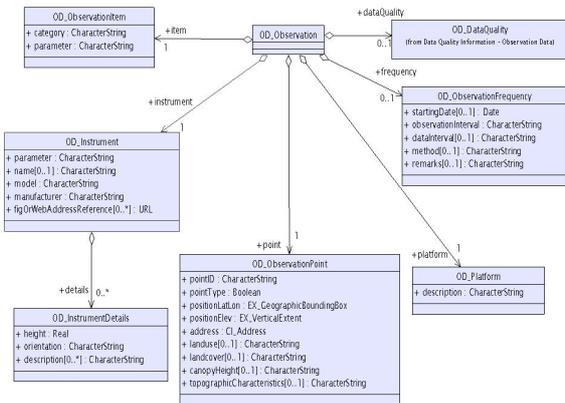


図4 河川観測データメタデータ 観測データ部分のUML

タが記述する内容の概要である。地上観測データは複数の観測項目を同時に観測しているため、複数の観測項目をグループ化したデータセットを単位として取り扱う必要がある。従って河川観測データメタデータはデータセット全体に関するメタデータと観測データ毎のメタデータを同時に記述する設計をとっている。図3はISO19115:2003を拡張した河川観測データメタデータの構成図である。色のついている部分が拡張した部分である。観測データの詳細、品質、品質管理と利用履歴情報を追加している。図4は観測データごとに記述する新たに作成したメタデータ項目の構成を表す。

先行研究[8]で述べたようにメタデータには“自動抽出可能なメタデータ”と“手入力メタデータ”がある。さらに手入力メタデータには、一度入力すると変更のない“静的なメタデータ”と、時間の経過によって変更が生じる“動的なメタデータ”がある。ほとんどのメタデータの内容は、データ作成者やデータ管理者が記述する必要がある。この作業は労力と時間を要する手入力メタデータ項目である。一方メタデータの長期管理には、動的なメタデータの管理が重要である。地球観測データのライフサイクルに合わせてメタデータの内容は蓄積されていくべきものであり、データの起源や系統を管理し、長期的なデータ管理を行うために必要な情報を動的メタデータ項目に蓄積する必要がある。

3. 現状のメタデータ登録システム

本節では、メタデータ登録システムの現状についていくつかのツールとともに紹介する。

3.1 メタデータ作成ツール

CSDGMやISO19115:2003形式のメタデータを作成するツールは多数存在する。CSDGMメタデータ形式のメタデータ作成ツールには、MetaLite^(注1)、CorpsMet^(注2)、NOAA ArcView Extension^(注3)、Xtme/tkme、DBFMeta^(注4)などがある。商用ではSMMS、CompuSult、ArcGIS and ArcCatalog、MetaS-

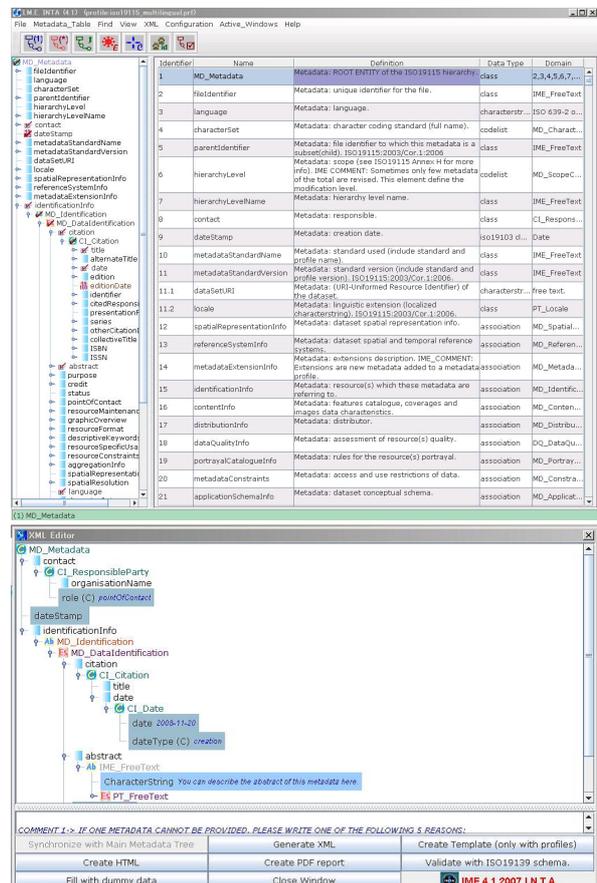


図5 ISO Metadata Editor 操作画面

tar などがある。一方ISO19115:2003形式のメタデータ作成ツールには、Enraemed^(注5)、MetaD^(注6)、CatMDEdit^(注7)、ISO Metadata Editor(IME)^(注8)などがあり、商用ではSMMS、CompuSult、ArcGISがある。これらのツールは、メタデータ作成者がオフラインでメタデータを作成するためのプログラムを配布するものが多い。作成されるメタデータファイルはCSDGMやISO19115:2003のメタデータ構造を満たすXMLファイルである。図5はIMEの操作画面である。IMEはISO19115のメタデータ項目を主要なメタデータ項目としたメタデータファイルを作成するツールである。このツールを利用してメタデータを入力するためには、メタデータの各項目の内容と作成したいメタデータ内容との関連付けを理解して操作する必要があり、メタデータに関する知識がない利用者が使うには使いにくいツールである。

3.2 メタデータ登録ツール

メタデータを作成するだけでなく、作成したメタデータを登録し管理するシステムがある。

docBuilder Metadata Authoring Tool

NASAのGCMD(Global Change Master Directory)が公開

(注1): <http://www.nsf.org.za/Metadata/metadataall.htm>

(注2): <http://corpsgeol.usace.army.mil/>

(注3): <http://www.csc.noaa.gov/metadata/download.html>

(注4): <http://geology.usgs.gov/tools/metadata/>

(注5): <http://clearinghouse4.fgdc.gov/enraemed/>

(注6): <http://www.geoportal-idec.net/geoportal/eng/inici.jsp?pag=metad&home=s>

(注7): <http://catmdedit.sourceforge.net/>

(注8): <http://www.crepad.rcanaria.es/metadata/en/index.en.htm>

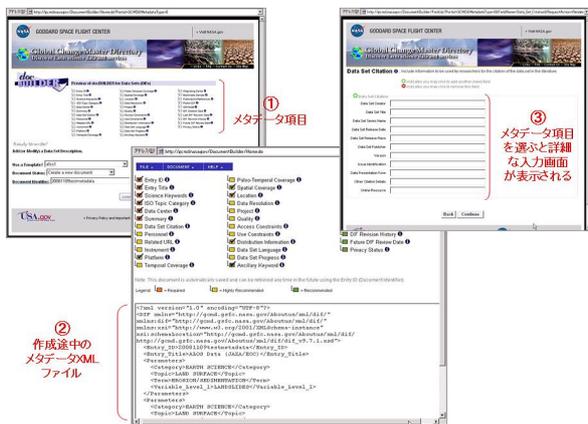


図 6 docBuilder 操作画面



図 7 GeoNetwork 操作画面

している docBUILDER^(注9)は GCMD が運用するメタデータ形式 DIF(Directory Interchange Format)のメタデータを作成し、登録することができるメタデータ作成・登録ツールである。図 6 はその操作画面である。①はメタデータ項目リストであり入力する項目を選択する。②は入力を反映するメタデータ XML 文書である。③はメタデータ項目の入力場面である。詳細なメタデータ項目と入力欄が表示される。docBuilder でのメタデータは Document Identifier で識別するため、入力内容の変更も Document Identifier を指定して行うことができる。さらにデータセットごとにメタデータのテンプレートを作成し、再利用することでメタデータ入力の負担を軽減させている。

GeoNetwork

GeoNetwork^(注10)は国際連合食糧援助機関 (FAO) が開発したオープンソースのメタデータカタログシステムである。メタデータの検索、編集と管理、さらに配布出版と共有機能を持つ。入力されたメタデータはメタデータ標準である ISO19115, FGDC の CSDGM, Dublin Core [1], [3] の XML に変換されデータベースに格納される。GeoNetwork サーバは GeoNetwork のノードとして登録することで他の GeoNetwork ノードとメタデータを交換する事ができるため、クリアリングハウスの構築

が容易にできる。図 7 はその操作画面である。メタデータの入力画面は default View, advanced View, XML View の 3 種類がある。default View では、責任者情報など必須項目を中心としたメタデータ入力項目が表示される。advanced View ではすべてのメタデータ項目が ISO19115 のパッケージで分類され、各パッケージごとに入力項目が表示される。XML View は、メタデータ XML ファイルそのものを入力したり編集することができる画面である。この GeoNetwork の入力画面について地球観測データを利用している関係者は、どのメタデータ項目が保有しているデータと関係するのかメタデータの項目の内容との対応付けが理解できず、このままでは利用が困難であるとコメントしている。

現状のメタデータ作成ツールは、どんなメタデータでも入力できるように汎用的な利用を目的としていて、メタデータ作成者が CSDGM や ISO19115:2003 の知識を持っていることが前提となっている。いくつかのツールでは、前もって入力する項目を絞り込んだテンプレートを作成しておくことができる。綿密にテンプレートを作成しておけば、メタデータ入力項目の選択を行なう負担が軽減できる。しかし、1 データセットに対してのメタデータ入力を 1 度で行なう設計であり、静的メタデータ、動的メタデータを区別することがないため、メタデータ入力者が本来入力する必要がない項目についても入力する必要がある。さらに我々が設計している河川流域メタデータは、ISO19115:2003 を拡張しているため、拡張部分のメタデータ項目を入れることができない。従ってこれらのツールをそのまま利用することはできない。

地球観測データを保有する観測者や観測機関の職員がメタデータを入力するため、メタデータについての知識を有しなくてもよいシステムが必要である。我々が必要とするメタデータ登録システムは次のような要件を満たすシステムである。

- (1) メタデータ入力者は Web インタフェースを利用してメタデータを入力する。クライアントにツールを配布しない。
- (2) サーバ側で入力されたメタデータをサーバ側が XML で管理する。メタデータ入力者は XML を意識する必要はない。
- (3) データの登録、データの品質検査、データ更新作業に伴うメタデータの更新を行なう。つまりメタデータの入力を独立させず、データアップロードからデータ品質管理を行なう一連の過程の途中で発生するメタデータも取り込みながら各データセットに対するメタデータを完成させていく。従って、メタデータの入力は複数の段階に分解して行なう。
- (4) 手入力メタデータ項目を極力減らす。
- (5) 責任者情報は、メタデータ項目として多数の箇所に入力する必要があるが、一度入力した責任者情報を再利用する。
- (6) すべてのメタデータに同一の値が入るメタデータ項目は、手入力せずあらかじめ設定しておく。
- (7) 自動生成可能なメタデータ項目は、入力画面に表示しない。
- (8) データの変更履歴、メタデータの更新履歴を管理する。

(注9): <http://gcmd.nasa.gov/DocumentBuilder/Home.do>

(注10): <http://geonetwork-opensource.org/>

4. 構築するデータアップロード、データ品質管理とメタデータ登録システム

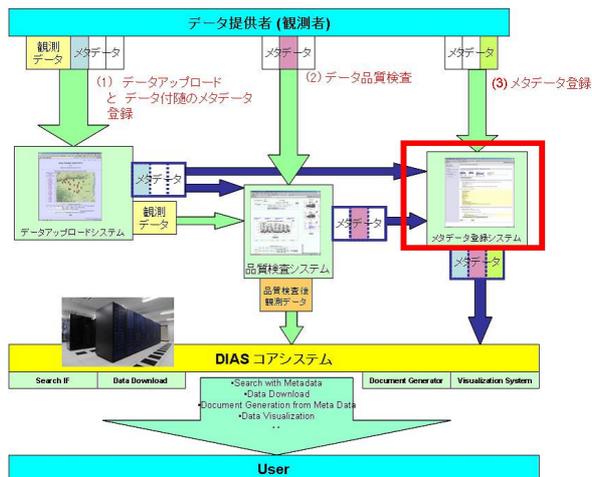


図 8 河川観測データアップロード、品質検査とメタデータ登録システム全体像

図 8 は、我々が開発している河川観測データアップロード、品質検査とメタデータ登録システムの全体像である。AWCI 参加国の担当者は (1) から (3) の作業を 3 つのシステムを通して行なう。いずれもインタフェースは Web インタフェースである。

(1) データアップロードとデータ付随のメタデータ登録
AWCI 参加国の担当者は保有のデータセットをデータアップロードシステムを通して Web インタフェースでアップロードする。その際、観測項目数、観測項目、単位、観測期間、観測間隔などを同時に登録する。これらの情報はこのデータセットに対するメタデータとして登録される。

(2) データ品質検査

アップロードされたデータセットをシステム側で品質検査を行なうデータベースに投入し、AWCI 参加国の担当者が Web インタフェースを通して品質検査を行なう。この品質検査システムは、すでに我々が他のデータセットで開発運用されているものを AWCI 河川観測データセットに適合させる [9], [10]。

(3) メタデータ登録

データセットに対するメタデータはデータアップロードシステム、データ品質検査システムによって自動的に更新される。さらに手入力が必要なメタデータ項目をメタデータ登録システムにおいて AWCI 参加国の担当者が入力する。

5. XForms を利用したメタデータ登録システム

2. で示したように、AWCI 河川流域メタデータは、メタデータの設計を UML で行い、メタデータ登録システムが扱うのはそのメタデータ構造を保有するメタデータインスタンスである。ISO19115:2003 の設計は ISO19139 として XML スキーマに変換され公開されている [5]。AWCI 河川流域メタデータについても同様に XML スキーマを生成し、そのスキーマが



図 9 UML から XForms XHTML ファイル作成までの過程

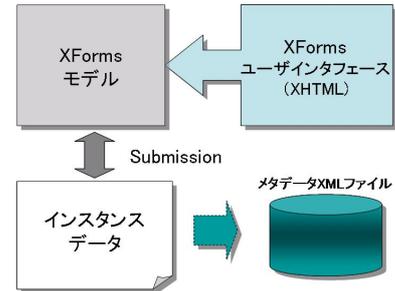


図 10 XForms の構成

らテンプレートとなる XML インスタンスを生成した。さらにテンプレート XML ファイルに基づき入力インタフェースを XForms [6], [7] で生成した。図 9 が UML から入力インタフェース XHTML 生成までの過程である。

(1) UML から XML スキーマを導出：The Eclipse Modeling Framework (EMF) を利用した。

(2) XML スキーマからテンプレートとなる XML ファイルの生成：Eclipse Web Standard Tools package(WTP) を利用した。

(3) テンプレート XML ファイルを出力するための XForms の生成：IBM XML Forms Generator を利用した。

ただし、生成した XHTML では、3. で紹介した他のツールと同様に、すべての入力項目の入力欄が列挙されるだけであり、そのままでは使えない。入力順、入力項目をひとつひとつ確認しながら現在のインタフェースを作成した。

5.1 XForms

XForms では今までの HTML フォームを、XForms モデル (M)、入力されるインスタンス (C)、そして XForms ユーザインタフェース (V)、の 3 つに分割することで、フォームの内容と表示を明確に分離した。XForms は構造化されている XML データの操作に加え、スキーマに対する妥当性検証、入力項目に対する制約条件の検証や計算を行うことができる。図 10 は XForms の MVC の関係の概要である。モデルとインスタンスをまず定義し、ユーザインタフェースについては XHTML を利用して記述している。入力終了時点で Submission を用いてメタデータ XML ファイルをデータベースに保存する。

XForms の実装は、クライアントで動作する単体のエンジン、ブラウザにプラグインして動作するエンジンがある。例えば前者には IBM Lotus Forms 3.0^(注11)、formsPlayer XForms Processor^(注12)がある。後者には Mozilla^(注13)、Mozilla XForms project^(注14)がある。これらは、クライアントに新たなアプリ

(注 11): <http://www-01.ibm.com/software/lotus/forms/>

(注 12): <http://www.formsplayer.com/>

(注 13): <http://sourceforge.net/projects/mozzie/>

(注 14): <http://www.mozilla.org/projects/xforms/>

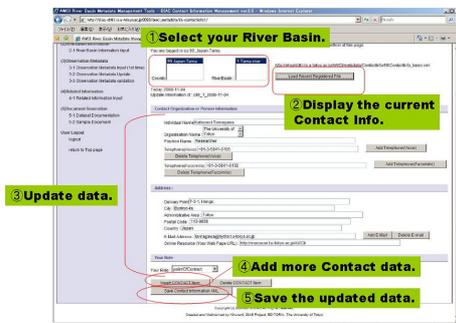


図 11 責任者+情報更新画面

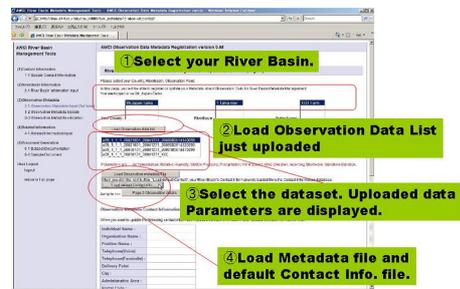


図 13 河川観測データメタデータ更新画面

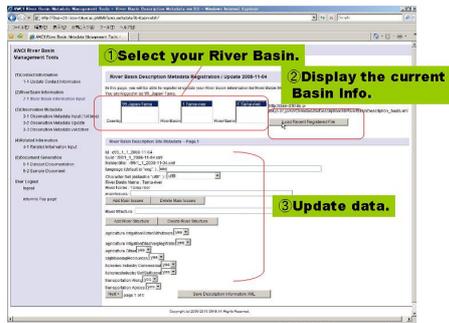


図 12 河川流域情報更新画面

面である。

6. まとめと今後の課題

本稿では AWCI 参加国が保有する地上観測データのメタデータ登録システムについて紹介した。地上観測データメタデータは、ISO19115:2003 を土台に拡張して設計されており、登録システムで作成したメタデータは XML データとして保存される。手入力メタデータの入力軽減をめざし、段階的に発生するメタデータを組み入れる方法をとっている。入力インターフェースはサーバサイド XForms エンジンを利用している。今後の課題は、実際に AWCI 参加国の担当者によるメタデータ入力を実施し利用者の意見を反映してインターフェースを改良すること。そして AWCI の参加国のネットワークの状況を考慮して現状のシステムの改良を行う。さらにドキュメント生成、進捗管理などのインターフェースを充実させることである。

文 献

- [1] DublinCore Metadata Initiative. *Dublin Core Metadata Element Set, Version 1.1*, 2008. <http://dublincore.org/documents/dces/>.
- [2] The Federal Geographic Data Committee (FGDC). *The Content Standard for Digital Geospatial Metadata*, 1998.
- [3] International Organization for Standardization. *ISO 15836:2003 Information and documentation - The Dublin Core metadata element set*.
- [4] International Organization for Standardization. *ISO 19115:2003 Geographic information - Metadata*, 2003.
- [5] International Organization for Standardization. *ISO/TS 19139:2007 Geographic information - Metadata - XML schema implementation*, 2007.
- [6] World Wide Web Consortium (W3C), <http://www.w3.org/TR/xforms/>. *XForms 1.0*, third edition, October 2007. W3C Recommendation.
- [7] World Wide Web Consortium (W3C), <http://www.w3.org/TR/xforms11/>. *XForms 1.1*, November 2007. W3C Candidate Recommendation.
- [8] 絹谷弘子, 生駒栄司, 高橋慧, 吉川正俊, 喜連川優. 地球観測データに対するメタデータ処理システムの設計. 電子情報通信学会第 19 回データ工学ワークショップ, 第 6 回日本データベース学会年次大会 (DEWS2008) 論文集, March 2008. C9-6.
- [9] 生駒栄司, 玉川勝徳, 小池俊雄, 喜連川優. 大規模地球環境観測データを対象としたデータクオリティコントロールシステムの構築とその有効性の検討. 日本データベース学会 Letters, Vol. Vol.4, No.1, pp. pp.57-60, 2005.
- [10] 生駒栄司, 谷口建司, 小池俊雄, 喜連川優. 大規模地球環境を対象としたデータクオリティコントロールシステムの構築. 第 16 回データ工学ワークショップ (DEWS2005) 講演論文集, 2005.

ケーションをインストールする必要がある。クライアントに新たな負担の必要がないサーバで動作する XForms エンジンとして Chiba project^(注15), Orbeon PresentationServer (OPS)^(注16)がある。サーバで XForms を XHTML に変換してからクライアントにデータを送付する。

サーバで動作する XForms エンジンを通じてクライアントに HTML によるデータ入力フォームを送付する方式がメタデータ登録ではクライアントに負担をかけないことからこの方式のエンジンを採用することにした。オープンソースで利用者の多い OPS を利用することにした。OPS は XForms1.0 と XForms1.1 の一部を実装している。OPS は XML 文書と XForms を土台とし、XML のキャプチャ、処理と表示に適し、Java や JavaScript などのコードを書く必要がない。XForms が MVC を分けて記述することで生成したい XML ファイルをモデルとして定義し、ビューを XHTML と XForms のタグで記述することができる。さらに他の XML ファイルを参照してモデルが持つ値を制御することができる。河川観測データメタデータではモデルは、図 3 で定義しているメタデータ XML である。インスタンスは、作成されるメタデータファイルのテンプレート XML, ログイン情報を保持する XML, ISO19115 の各種コードリスト XML, 流域情報メタデータ XML, 責任者情報メタデータ XML と入力フォームで選んだ選択肢の値を持つ XML などである。

5.2 メタデータ登録インターフェース

図 11 から図 13 が作成したメタデータ登録システムの表示画

(注15): <http://sourceforge.net/projects/chiba/>

(注16): <http://www.orbeon.com/>