

# 地球環境データ統融合利用促進に向けた ドキュメントセントリックメタデータ作成と検索

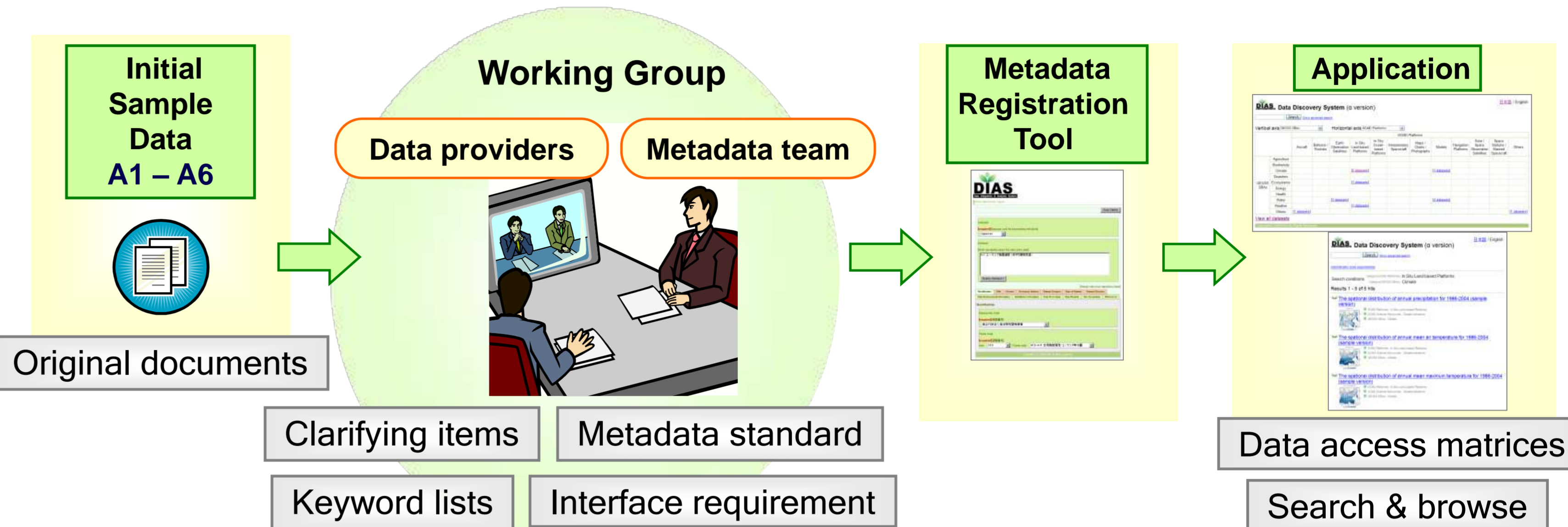
Registration and retrieval of document centric metadata  
for accelerating integrated use of earth environmental data

絹谷弘子<sup>1</sup>, 清水敏之<sup>2</sup>, 吉川正俊<sup>2</sup>, 小野雅史<sup>1</sup>, 福田和代<sup>3</sup>, 市野美夏<sup>3</sup>, 川本温子<sup>3</sup>

<sup>1</sup>東京大学地球観測データ統融合連携研究機構, <sup>2</sup>京都大学情報学研究科, <sup>3</sup>海洋研究開発機構

地球観測データおよびモデル出力データの可用性、流通性を高めるためには、メタデータの整備が不可欠である。多種多様な地球科学データに対してメタデータを管理することで、データに対する理解を支援し、データの統融合利用が促進する。我々はメタデータを作成するためのシステム開発をデータ提供者とシステム開発者の協働で行った。データ提供者はデータを説明するためのドキュメントを記述する機会が多いため、ドキュメント記述と同時にメタデータが作成される仕組みを考えた。まず、データ提供者が作成したドキュメントから、共通の情報を抽出し、メタデータ入力項目の検討を行った。それを基に、Webアプリケーションとしてメタデータ作成ツールを開発した。PDF、HTML形式でドキュメントが作成され、ISO19139に基づくXML形式でメタデータが作成される。さらに、メタデータを用いたデータ検索システムとして、キーワード検索やデータ俯瞰のためのマトリクス表示等を実装した。

Collaborative development activities with data providers (JAMSTEC) and system developers (Universities)



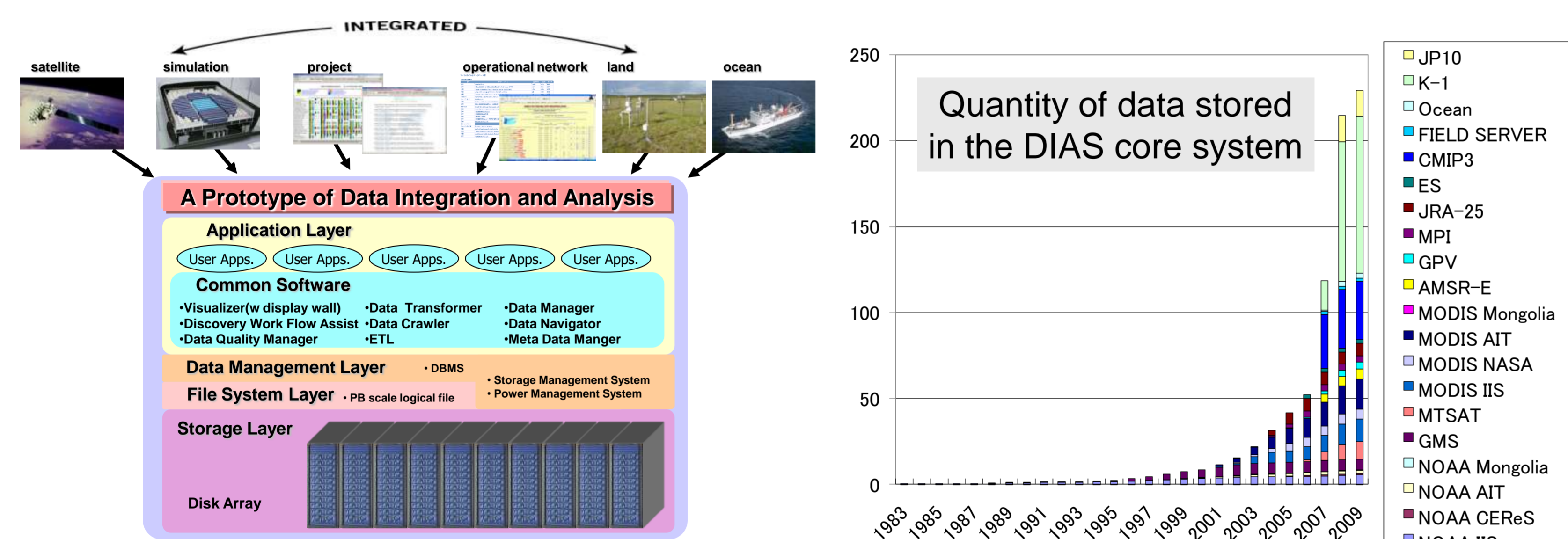
## Data Integration & Analysis System (DIAS)

DIAS (Data integration and Analysis System) では、多種多様な地球観測データやデータプロダクトをDIASコアシステムに格納し、研究成果の一部はWebサイトから公開されている。

<http://www.jamstec.go.jp/e/medid/dias>

各データプロダクトを気候、水循環、農業、生態系などの専門分野を超えて活用することで、DIASの共有財産であるデータプロダクトが生み出す価値を高めることができる。

<http://www.editoria.u-tokyo.ac.jp/dias/>

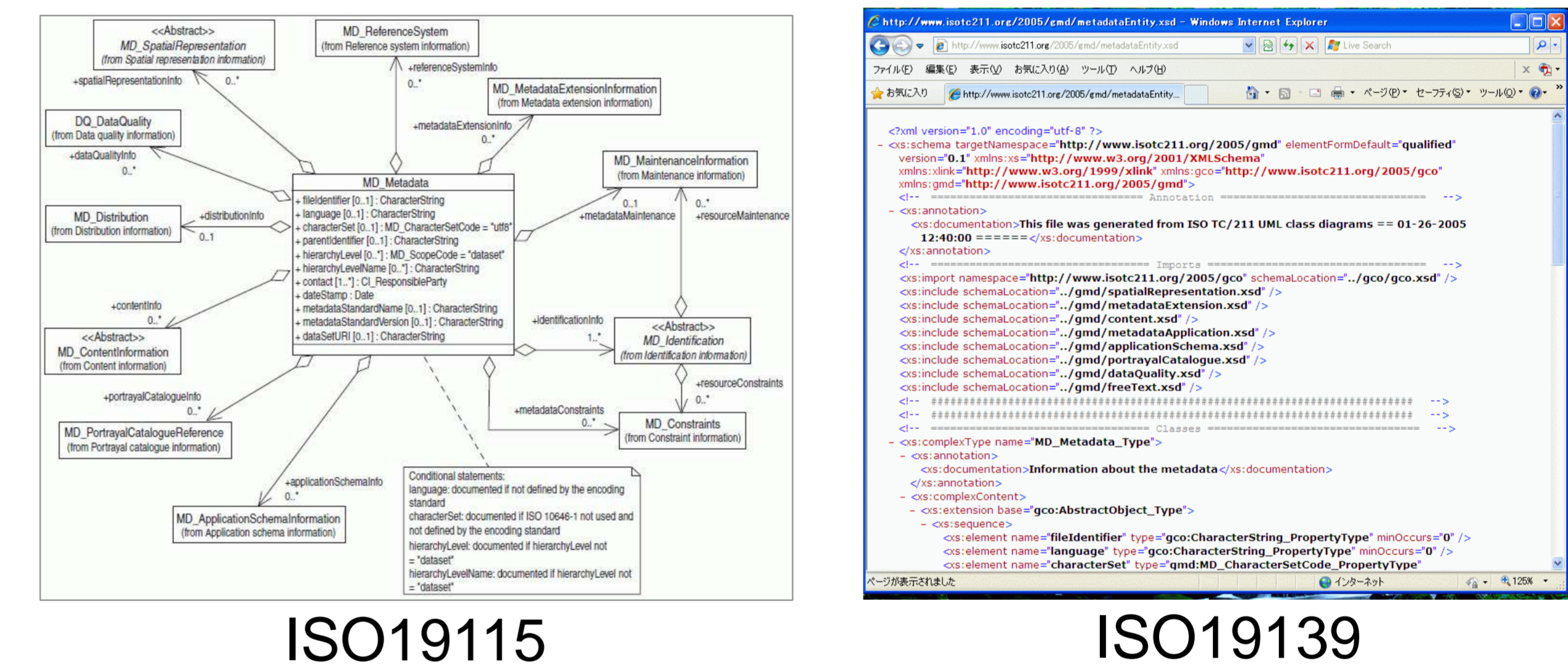


## ISO19115/19139 メタデータ

ISO (the International Organization for Standardization) は国際規格を策定するための組織である。

ISO 19115:2003 は地理情報の分野におけるメタデータの論理的な仕様をUML(Unified Modeling Language) で定義したものであり、ISO 19139:2007 はXML(Extensible Markup Language) を用いたISO 19115の実装の一つである。

ISO19115/19139の要素としては、連絡者、データの制約、空間、時間、データの品質などがある。



地球科学データを用いた研究活動を支えるためには、良質なメタデータを作成し、維持管理を行なうことが重要である。データを理解しているデータプロダクト作成者がメタデータを作成することが望ましいが、既存のメタデータ作成ツールはメタデータに関する専門知識を必要とする場合が多く、データプロダクト作成者によるメタデータ作成が困難となっている。

我々はデータプロダクト作成者が作成するデータセットを説明するドキュメントの生成支援を行い、並行してデータセットに関するメタデータを作成できるツールを作成した。このツールはWebアプリケーションであり、プロダクト作成者が新たなソフトウェアのインストールする必要がない。インターネットを介して入力ボックスに必要な情報を順次入力するだけでデータセットに関するドキュメントが自動的にPDF、HTML形式で作成され、さらにISO19139標準形式のメタデータXMLファイルが作成される。作成されたドキュメントはすぐに利用でき、メタデータ作成に新たな負担をかけない点が特長である。作成されたメタデータはDIASプロジェクトのメタデータデータベースが維持管理することで、DIASコアシステムと研究活動との一体的な運営を行なうことができる。

さらに、我々はメタデータを利用したデータ俯瞰、検索、データダウンロードのためのシステムを開発中である。キーワード検索、場所検索、期間検索の機能により、データセットを絞り込むことが可能であり、また、マトリクス表示により、容易にDIASに登録されているデータセット全体を俯瞰することが可能である。

